(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 July 2002 (04.07.2002)

PCT

(10) International Publication Number
WO 02/052047 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US01/50242

(22) International Filing Date:
21 December 2001 (21.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/258,070 22 December 2000 (22.12.2000) US

(71) Applicant *(for all designated States except US)*: **MERGEN LTD.** [US/US]; 14826 Wicks Blvd, San Leandro, CA 94577 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **YU, Zailin** [US/US]; 2374 Belvedere Ave., San Leandro, CA 94577 (US). **PENG, Zaoyuan** [US/US]; 865 San Jude St., Palo Alto, CA 94306 (US). **WILEY, Andrew** [US/US]; 1093 Rose Drive, Napa, CA 94558 (US). **HU, Qianjin** [CN/US]; 20499 Crow Creek Rd., Castro Valley, CA 94552 (US).

(74) Agents: **BASE, Shantanu** et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/052047 A2

(54) Title: METHODS FOR IDENTIFYING G-PROTEIN COUPLED RECEPTORS ASSOCIATED WITH DISEASES

(57) Abstract: The present invention describes a method for isolating novel G-protein coupled receptors (GPCRs) whose expression is associated with disease states. The invention provides polynucleotide and polypeptide sequences whose expression is associated with neurodegenerative diseases and cancer. The invention also provides evidence of novel G-protein coupled receptors (GPCRs) related polynucleotide sequences whose expression level is associated with Parkinson's disease, Alzheimer's disease and leukemia. The present invention provides methods for identifying GPCR polypeptides associated with diseases and provides diagnostic compositions and methods for the detection of diseases.

METHODS FOR IDENTIFYING G-PROTEIN COUPLED RECEPTORS ASSOCIATED WITH DISEASES

INVENTORS

Zaoyuan PENG; Zailin YU; Andrew WILEY and Qianjin HU

TECHNICAL FIELD OF THE INVENTION

[0001] The present invention relates to the field of detection of gene expression. In particular, the present invention relates to methods for identifying novel genes whose expression is associated with disease states. More particularly, the present invention relates to novel GPCR polynucleotides and polypeptides associated with CNS diseases and cancer.

BACKGROUND

[0002] G-protein linked cell receptors comprise the largest superfamily of mammalian proteins with more than 1,000 different members identified to date. (Gether, U. Endocr. Rev. 21(1):90-113 (Feb 2000)). GPCRs are found in a very wide range of species, and are often involved in signal transduction across cell membranes. GPCR sequences include a characteristic seven hydrophobic transmembrane domains which span the plasma membrane and form a bundle of antiparallel alpha helices, commonly identified as seven transmembrane (7-TM) helices. The transmembrane domains account for structural and functional features of the receptor. Ligands which activate the various members of the GPCR family include an enormous variety of molecules such as amines, amino acids and peptides as well as several small hydrophobic molecules. Usually the 7-TM bundle of helices forms a binding pocket. When the binding site needs to accommodate bulky molecules, the extracellular N-terminal segment or one or more of the three extracellular loops also participate in binding and in subsequent induction of conformational change in intracellular portions of the receptor. The activated receptor, in turn, interacts with an

intracellular heterotrimeric G-protein complex which mediates further intracellular signaling activities, such as interactions with guanine nucleotide binding (G) proteins and the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate or ion channel proteins (Baldwin, J. M. (1994) Curr. Opin. Cell Biol. 6:180-190). GPCRs may also function via other activation mechanisms such as, photon- activation of opsins and self-activation of thrombin receptor after cleavage of its N-terminus by thrombin. (*see*, Watson, S. and S. Arkinstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego, Calif.).

[0003] Among membrane-bound receptors, the 7-TM receptors are the most abundant, comprising up to 1% of the genome. They transduce signals in response to a large number of physiologically important molecules, including photons, organic odorants, nucleotides, nucleosides, biogenic amines, sugars, peptides, lipids, proteins and viruses. They also catalyze GDP/GTP nucleotide exchange on heterotrimeric G-proteins and therefore are also referred to as G-protein-coupled receptors (GPCRs). To date, greater than 60% of the 400 or so drug targets are GPCRs and these receptors play a key role in homeostatic regulation in many tissues.

[0004] The amino-terminus of a GPCR is extracellular, of variable length and often glycosylated, while the carboxy-terminus is cytoplasmic and generally phosphorylated. Extracellular loops of GPCRs alternate with intracellular loops and link the transmembrane domains. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. GPCRs range in size from under 400 to over 1000 amino acids (Coughlin, S. R. (1994) Curr. Opin. Cell Biol. 6:191-197).

[0005] It is generally believed that members of the 7-TM helix G protein-coupled receptor family pair up with their own kind but do not bind to other family members (Hebert, T. E. and Bouvier, M., *Biochem. Cell Biol.* **76**, 1 (1998)). However, recent studies provide evidence that GPCRs can pair up with even rather distantly related relatives to form heterodimeric receptors with distinct properties (Rocheville, M. *et al.*, *Science* **288**, 154-157 (2000)).

[0006] A highly conserved D-R-Y motif comprising the amino acids Asp-Arg-Tyr is also characteristic of GPCRs (Acharya, S., and Karnik, S. S. *J. Biol. Chem.* 271:25406-25411 (1996)) and is involved in constitutive activity and structural stability of GPCRs. (*see,* Alewijnse, A.E. et al., *Mol Pharmacol* 57(5):890-898 (2000)).

[0007] GPCRs respond to a diverse array of ligands including lipid analogs, amino acids and their derivatives, peptides, cytokines, and specialized stimuli such as light, taste, and odor. GPCRs function in physiological processes including vision (the rhodopsins), smell (the olfactory receptors), neurotransmission (muscarinic acetylcholine, dopamine, and adrenergic receptors), and hormonal response (luteinizing hormone and thyroid-stimulating hormone receptors).

[0008] GPCR mutations, affecting both loss-of-function and activation, have been associated with numerous human diseases (Coughlin, supra). For instance, retinitis pigmentosa may arise from either loss-of-function or activating mutations in the rhodopsin gene. Somatic activating mutations in the thyrotropin receptor cause hyperfunctioning thyroid adenomas (Parma, J. et al. (1993) Nature 365:649-651). Parma et al. also suggest that certain GPCRs susceptible to constitutive activation may be proto-oncogenes.

[0009] Neuropeptide Y (NPY), pancreatic polypeptide (PP), and peptide YY (PYY) are structurally related peptides found in higher vertebrates. NPY is produced in the central and peripheral nervous systems. NPY plays a role in the stimulation of food intake, anxiety, facilitation of learning and memory, and regulation of the neuroendocrine and cardiovascular systems. NPY also stimulates vascular smooth muscle contraction, modulates hormone secretion, and has been implicated in the pathophysiology of hypertension, congestive heart failure, affective disorders and appetite regulation (Watson, S. and S. Arkinstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego Calif., pp. 194-198). PP is produced in endocrine cells in the pancreas and inhibits pancreatic secretion, gall bladder contraction, and gut motility. PYY is produced in endocrine cells of the pancreas and large intestine. PYY has actions similar to those of PP, and in addition redistributes blood flow in gut vessels. Both PP and PYY

are released into the circulation in response to food intake. These structurally related peptides accomplish their varied biological functions through interaction with distinct GPCR subtypes. Several receptor subtypes have been defined by their ability to bind NPY, PYY, PP, and derivatives of these peptides. At least five distinct receptor subtypes have been characterized to date (Weinberg, D. H. et al. (1996) J. Biol. Chem. 271:16435-16438).

[0010] Cholecystokinin (CCK) is a bioactive peptide which is present throughout the digestive tract, and is also found in smooth muscle tissues such as bladder and uterus, in secretory glands such as exocrine pancreas, and in the brain. The major physiological actions of CCK are gall bladder contraction, pancreatic enzyme secretion, and regulation of secretion/absorption in the gastrointestinal tract. CCK receptors are GPCRs found in peripheral tissues including pancreas, stomach, intestine and gall bladder, and in limited amounts in the brain. CCK receptors mediate pancreatic acinar secretion and gallbladder contraction (De Weerth, A. et al. (1993) Am. J. Physiol. 265:G1116-G1121). The CCK-A receptor has been implicated in the pathogenesis of schizophrenia, Parkinson's disease, drug addiction and feeding disorders (Watson and Arkinstall, supra, pp. 89-95).

SUMMARY OF THE INVENTION

[0011] The present invention describes a method for isolating novel G-protein coupled receptors (GPCRs) whose expression is associated with diseases. The present invention relates to novel GPCR polypeptides associated with CNS diseases and cancer and antibodies specific for the GPCR polypeptides in the detection, diagnosis, prevention and/or treatment of the GPCR expression-related diseases.

[0012] The present invention relates to materials and methods for diagnosing, preventing and/or treating GPCR expression-related diseases in subjects. The present invention is based, in part, on the discovery of differential levels of expression of GPCR-related polynucleotides in tissue samples of Parkinson's disease, Alzheimer's disease and leukemia patients, as compared to normal tissue samples. The present invention provides polynucleotides, as well as their corresponding GPCR gene products, that are present at

elevated levels in these disease states. These polynucleotides (or fragments thereof) and polypeptides (or antigenic fragments thereof), and antibodies that bind such polypeptides, are useful in a variety of diagnostic, prophylactic and therapeutic methods.

[0013] The present invention provides a novel GPCR polynucleotide having the sequence illustrated in Figure 1B (SEQ ID NO 1) for use in the detection, diagnosis, prevention and treatment of Alzheimer's diseases and leukemia.

[0014] The present invention provides a novel GPCR polynucleotide having the sequence illustrated in Figure 2B (SEQ ID NO 2) for use in the detection, diagnosis, prevention and treatment of Parkinson's and Alzheimer's diseases and leukemia.

[0015] The present invention provides a novel GPCR polynucleotide having the sequence illustrated in Figure 3B (SEQ ID NO 3) for use in the detection, diagnosis, prevention and treatment of Parkinson's and Alzheimer's diseases and leukemia.

[0016] The present invention provides a novel GPCR polynucleotide having the sequence illustrated in Figure 4B (SEQ ID NO 4) for use in the detection, diagnosis, prevention and treatment of Parkinson's and Alzheimer's diseases and leukemia.

[0017] The present invention provides a novel GPCR polynucleotide having the sequence illustrated in Figure 5B (SEQ ID NO 5) for use in the detection, diagnosis, prevention and treatment of leukemia.

[0018] A method is provided for identifying a gene whose expression level is associated with a disease state, the method comprising: identifying at least one gene having a nucleic acid sequence encoding a protein comprising a physical characteristic; selecting a polynucleotide sequence from the nucleic acid sequence, wherein the polynucleotide sequence is specific for a protein comprising the physical characteristic; detecting a level of expression of the polynucleotide sequence or a complement thereof in a diseased tissue sample; detecting a level of expression of the polynucleotide sequence or a complement thereof in a normal tissue sample; and comparing the level of expression of the polynucleotide sequence or a complement in the diseased tissue sample to a level of expression of the gene in the control tissue sample, wherein an altered level of expression

of the polynucleotide sequence or a complement in the diseased tissue sample correlates with the disease state.

[0019] The physical characteristic may comprises seven transmembrane (7-TM) domains, or an Asp-Arg-Tyr (DRY) motif, or a signal peptide sequence characteristic of a secreted protein, or a signal peptide sequence characteristic of a mitochondrial protein, or an amino acid sequence characteristic of a structural feature of the protein, or an amino acid sequence characteristic of a function of the protein.

[0020] The identification of the gene may comprise searching a nucleic acid sequence database for nucleic acid sequences which encode a protein comprising the physical characteristic. The nucleic acid sequence database is optionally an electronic library. The gene is optionally identified using a search algorithm. The identification of the gene may further comprise selecting at least one gene whose expression is known to correlate with a disease state. In one embodiment, the gene is a novel G-protein linked cell receptor type gene

[0021] In one embodiment of the method, the detection of the level of expression of the polynucleotide sequence comprises: selecting at least one isolated oligonucleotide comprising the polynucleotide sequence or a fragment thereof; contacting the oligonucleotide with a nucleic acid preparation from the tissue sample; and detecting a level of expression of the polynucleotide sequence by detecting an amount of hybridization of the nucleic acid preparation to the oligonucleotide under stringent conditions.

[0022] In one embodiment, the oligonucleotide is attached to a solid support. In another embodiment, the solid support is a microarray.

[0023] In one embodiment of the method, the selection of the polynucleotide sequence comprises determining at least one of a set of factors comprising (i) a redundancy of the sequence, (ii) an efficiency of hybridization to a complementary sequence, and (iii) a likelihood of the polynucleotide sequence comprising an intron.

[0024] The nucleic acid preparation from the tissue sample may comprise a detectable label. The detectable label is optionally selected from the group consisting of a

fluorescent label, an enzymatic label, a chemiluminescent label, a colorimetric label, and a radioactive label. The nucleic acid preparation from the tissue sample is optionally amplified before detection. In one embodiment, the amplification is a conducted by a polymerase chain reaction (PCR). In one embodiment, the amplification is a conducted by a quantitative polymerase chain reaction (QPCR).

[0025] In one embodiment of a method according to the invention, comparing the level of expression of the gene comprises: providing at least one isolated oligonucleotide comprising the polynucleotide sequence or a fragment thereof; contacting the oligonucleotide with an amount of nucleic acid preparation from a disease tissue sample; contacting the oligonucleotide with an equal amount of nucleic acid preparation from a normal tissue sample; and comparing the level of expression of the polynucleotide sequence in the tissue samples by detecting an amount of hybridization of each nucleic acid preparation to the oligonucleotide under stringent conditions.

[0026] In some embodiments, the polynucleotide is attached to a solid support, wherein the solid support is optionally microarray. The nucleic acid preparation is optionally an RNA preparation, wherein the RNA preparation may be further processed to generate a labeled nucleic acid probe. The labeled nucleic acid probe may comprise a label coupled to the probe, wherein the label is selected from the group consisting of a biotin, an avidin, a streptavidin, an antibody, an antigen, a peptide, a fluorescent label, an enzymatic label, a chemiluminescent label, a colorimetric label, and a radioactive label.

[0027] The invention also provides a method for detecting in a cell an expression of a gene whose expression level is associated with a disease state, the method comprising: cloning a polynucleotide fragment comprising a sequence of the cloned gene in an expression vector; and detecting a corresponding protein in a cell transformed with the vector comprising the cloned fragment. In some embodiments the protein is detected by an antibody, or a monoclonal antibody.

[0028] The invention further provides a method for preparing an antibody specific for a polypeptide product of a gene whose expression level is associated with a disease state, the method comprising: cloning a polynucleotide fragment comprising a sequence of the

cloned gene in an expression vector; isolating a polypeptide expressed by the vector, wherein the polypeptide comprises an amino acid sequence corresponding to the cloned polynucleotide; immunizing an animal with the isolated polypeptide; and isolating anti-peptide antibodies specific for the isolated polypeptide from the immunized animal.

[0029] The invention provides an isolated novel polynucleotide comprising a gene whose expression level is associated with a disease state, the polynucleotide comprising a nucleic acid sequence encoding a protein comprising a physical characteristic, wherein the polynucleotide or a fragment thereof is differentially expressed in a diseased tissue sample as compared to a normal tissue sample.

[0030] The isolated polynucleotide may comprise a nucleic acid sequence which encodes a protein comprising at least one of the characteristics of: (a) seven transmembrane (7-TM) domains, (b) an amino acid sequence comprising an Asp-Arg-Tyr (DRY) motif and (c) a signal peptide. The isolated polynucleotide may comprise a detectable label. The isolated polynucleotide, or fragment thereof, is optionally attached to a solid support. The isolated polynucleotide may be single or double stranded.

[0031] The invention further provides a host cell, an array, a composition comprising a test cell sample and an electronic library, each comprising at least one isolated polynucleotide of the invention.

[0032] The present invention provides a method for the detection and diagnosis of diseases, the method comprising detecting an altered level of expression of the GPCR polynucleotides comprising sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5). In one aspect of the invention, a diagnosis of Parkinson's disease, Alzheimer's disease and/or leukemia correlates to an altered level of expression of the polynucleotides comprising sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) in disease tissue samples. In one aspect of the invention, a diagnosis of Parkinson's and/or Alzheimer's disease correlates to a lower level of expression of the polynucleotides comprising sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) in patient tissue samples. In another aspect of the invention, a diagnosis of leukemia disease

correlates to an elevated level of expression of the polynucleotides comprising sequences shown in Figures 1B, 2B, 3B, 4B and 5B in patient tissue samples.

[0033] The present invention also provides polynucleotides comprising a segment of at least 10 nucleotides in length which corresponds identically to a portion of the sequences shown in Figures 1B, 2B, 3B, 4B and 5B and whose expression correlates to a state of disease. In further embodiments, the segment of the polynucleotide sequence is up to at least 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000, or 1500 nucleotides in length.

[0034] In other embodiments, a polynucleotide of the present invention comprises a detectable label, and/or is attached to a solid support. In other embodiments, a polynucleotide of the present invention is single stranded and in yet other embodiments, is double stranded. The present invention also encompasses host cells comprising an isolated polynucleotide of the present invention.

[0035] In a further aspect, the present invention provides GPCR polypeptides comprising polypeptides having sequences shown in Figures 1C, 2C, 3C, 4C and 5C, or fragment thereof, whose presence and /or expression levels correlate to a state of central nervous system (CNS) disease, Parkinson's disease, Alzheimer's disease and leukemia. In one embodiment, the invention provides GPCR polypeptides having sequences shown in Figures 1C, 2C, 3C, 4C and 5C, or fragment thereof, whose presence and /or expression levels correlate to a state of Parkinson's and/or Alzheimer's disease and/or leukemia. In an additional embodiment, the polypeptide comprises a fragment that includes an epitope of the amino acid sequence shown in Figures 1C, 2C, 3C, 4C and 5C. In additional embodiments, the polypeptide or fragment thereof, is attached to a solid support.

[0036] The present invention also provides isolated antibodies or antigen binding fragments thereof, that bind to a polypeptide of the present invention. The present invention also provides isolated antibodies or antigen binding fragments thereof, that bind to a polypeptide encoded by a polynucleotide of the present invention. The present invention also provides isolated antibodies that bind to a polypeptide of the invention, or antigen binding fragment thereof, encoded by a polynucleotide made by the method

comprising the following steps i)immunizing a host animal with a composition comprising said polypeptide of the present invention, or antigen binding fragment thereof, and ii) collecting cells from said host expressing antibodies against the antigen or antigen binding fragment thereof.

[0037] The present invention also provides isolated antibodies that bind to a polypeptide, or antigen binding fragment thereof, encoded by a polynucleotide of the present invention made by the method comprising the following steps: providing a cell line producing an antibody, wherein said antibody binds to a polypeptide of the present invention, or antigen binding fragment thereof, encoded by a polynucleotide of the present invention and culturing said cell line under conditions wherein said antibodies are produced. In additional embodiments, the antibodies are collected and monoclonal antibodies are produced using the collected host cells or genetic material derived from the collected host cells. In additional embodiments, the antibody is a polyclonal antibody. In a further embodiment, the antibody is attached to a solid surface or further comprises a detectable label.

[0038] The present invention also discloses a method for diagnosing a disease associated with a G-protein coupled receptor (GPCR) expression in a test tissue sample, the method comprising: detecting a level of expression of a polynucleotide comprising a GPCR DNA sequence or a fragment thereof; and comparing the level of expression of the polynucleotide in the test tissue sample to a level of expression in a control tissue sample, wherein an altered level of expression of the polynucleotide in the test tissue sample is indicative of a disease state.

[0039] The disease state may be selected from the group consisting of a central nervous system (CNS) disease, Parkinson's disease, Alzheimer's disease and leukemia.

[0040] The polynucleotide may comprise a GPCR DNA fragment at least 10 contiguous nucleotides in length. In some embodiments, the polynucleotide comprises: (a) a polynucleotide having sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement; (b) a fragment of the polynucleotide comprising the sequence shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement, wherein

the fragment is at least 10 nucleotides in length; or (c) a polynucleotide that selectively hybridizes to the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) or the fragment in (b).

[0041] The altered level of expression may be an elevated or a lowered level of expression.

[0042] The detecting optionally comprises measuring the level of an RNA comprising a sequence complementary to the polynucleotide, or its complement. The polynucleotide may comprise a detectable label. The detectable label is selected from the group consisting of a fluorescent label, an enzymatic label, a chemiluminescent label, a colorimetric label, and a radioactive label. The polynucleotide is optionally attached to a solid support, or a microarray. The polynucleotide may be single or double stranded. The polynucleotide may be amplified before detection. The amplification is optionally conducted by a conducted by polymerase chain reaction (PCR).

[0043] A lowered level of expression of the polynucleotide in the test tissue sample may be indicative of a CNS disease, a Parkinson's disease or a Alzheimer's disease. An elevated level of expression of the polynucleotide in the test tissue sample may be indicative of a leukemia disease.

[0044] The invention also provides methods for diagnosing a disease associated with a G-protein coupled receptor (GPCR) expression in a test tissue sample, the method comprising: detecting a level of expression of a polypeptide comprising sequences shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10), or a fragment thereof; and comparing the level of expression of the polypeptide in the test tissue sample to a level of expression in a control tissue sample, wherein an altered level of expression of the polypeptide in the test tissue sample is indicative of a disease state.

[0045] The disease state may be selected from the group consisting of a CNS disease, Parkinson's disease, Alzheimer's disease and leukemia. The polypeptide may comprise a fragment comprising at least one epitope of the amino acid sequences shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10). The polypeptide or fragment thereof, may be attached to a solid support or an array.

[0046] The invention further includes an isolated polynucleotide comprising: (a) a polynucleotide having sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement; (b) a fragment of the polynucleotide having the sequence shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement, wherein the fragment is at least 10 nucleotides in length; or (c) a polynucleotide that selectively hybridizes to the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) or the fragment in (b), wherein expression of the isolated polynucleotide correlates to a state of disease.

[0047] The isolated polynucleotide may comprise a segment of up to at least 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000 or 1500 nucleotides in length which corresponds identically to a portion of the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5).The isolated polynucleotide may further comprise a detectable label. The isolated polynucleotide may be attached to a solid support, be single or double stranded.

[0048] The invention further provides a host cell, an array comprising at least two polynucleotides according, a composition, comprising a test cell sample and an isolated polynucleotide, a kit for diagnosing a GPCR-related disease in a test sample, and an electronic library comprising at least one isolated polynucleotide comprising the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5).

[0049] The invention provides an isolated polypeptide encoded within a GPCR open reading frame, or a fragment thereof, whose expression levels in a tissue correlates to a disease state of the tissue. In some embodiments, the isolated polypeptide comprises the amino acid sequence encoded by the polypeptides shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10) or a fragment thereof. The disease state may be selected from the group consisting of central nervous system (CNS) disease, Parkinson's disease, Alzheimer's disease and leukemia. The polypeptide optionally comprises a fragment that includes an antigenic epitope comprising the amino acid sequence shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10).

[0050] The invention also provides an isolated antibody, or antigen binding fragments thereof, that bind to the polypeptide according to the invention. The antibody may be a polyclonal or monoclonal antibody, and may be attached to a solid surface and may comprise a detectable label.

[0051] The invention provides a method for producing an antibody, or antigen binding fragments thereof, that bind to a polypeptide according to the invention, the method comprising: (a) immunizing a host animal with a composition comprising the polypeptide, or fragment thereof; and (b) collecting one or more cells from said host, wherein said cells produce antibodies against the polypeptide or fragment thereof. In one embodiment the method further comprises: (c) producing monoclonal antibodies using the collected host cells or genetic material derived from the collected host cells.

[0052] The invention provides a method for screening a drug candidate for activity against a GPCR-related disease, the method comprising: (a) contacting a tissue sample derived from a cell associated with an altered GPCR-related disease with a drug candidate; (b) monitoring the expression in the tissue sample of a polynucleotide comprising the sequence shown in Figs. 1B, 2B, 3B, 4B or 5B (SEQ ID NOS 1-5), or a complement thereof; and (c) determining the efficacy of the drug candidate.

[0053] The invention provides a method for detecting a disease associated with expression of a GPCR polypeptide in a test cell sample, the method comprising: (a) detecting a level of expression of at least one polypeptide, or a fragment thereof, according to the invention; and (b) comparing said level of expression of the polypeptide in the test sample with a level of expression of polypeptide in the control cell sample, wherein an altered level of expression of the polypeptide in the test cell sample relative to the level of expression of the polypeptide in the control cell sample is indicative of the presence of the disease in the test cell sample.

[0054] The invention provides a method for detecting a disease associated with the presence of an anti-GPCR antibody in a test cell sample, the method comprising: (a) detecting a level of an antibody against a polypeptide, or fragment thereof, according to the invention; and (b) comparing said level of said antibody in the test sample with a level

of said antibody in the control cell sample, wherein an altered level of antibody in said test cell sample relative to the level of antibody in the control cell sample is indicative of the presence of the disease in the test cell sample.

[0055] The invention provides a method for stimulating an immune response in a human against cells that express GPCR polypeptides, the method comprising administering to a human an immunogenic amount of: (a) a polypeptide, comprising an amino acid sequence of a GPCR polypeptide according to the invention; or (b) a mutein or variant of a polypeptide comprising the amino acid sequence of a GPCR polypeptide according to the invention.

[0056] The present invention also provides compositions comprising a test cell sample and an isolated polynucleotide of the present invention. The present invention further provides methods for detecting disease associated with expression of a polypeptide in a test cell sample, comprising the steps of: i) detecting a level of expression of at least one polypeptide of the present invention, or a fragment thereof and ii) comparing said level of expression of the polypeptide in the test sample with a level of expression of polypeptide in the control cell sample, wherein an altered level of expression of the polypeptide in the test cell sample relative to the level of expression of the polypeptide in the control cell sample is indicative of the presence of the disease in the test cell sample. The present invention also provides methods for detecting disease associated with the presence of an antibody in a test cell sample, comprising the steps of: i) detecting a level of an antibody of the present invention, and ii) comparing said level of said antibody in the test sample with a level of said antibody in the control cell sample, wherein an altered level of antibody in said test cell sample relative to the level of antibody in the control cell sample is indicative of the presence of disease in the test cell sample.

BRIEF DESCRIPTION OF THE FIGURES

[0057] Figure 1A illustrates the expression levels in Parkinson's and Alzheimer's brain tissue and white blood cells of a leukemia patient of the GPCR gi6863021_GS_nt7 having the polynucleotide sequence shown in Fig. 1B (SEQ ID NO 1), the amino acid

sequence of shown in Fig. 1C (SEQ ID NO 6) and the Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 1D.

[0058] Figure 2A illustrates the expression levels in Parkinson's and Alzheimer's brain tissue and white blood cells of a leukemia patient of the GPCR gi6453999_GS_nt6 having the polynucleotide sequence shown in Fig. 2B (SEQ ID NO 2), the amino acid sequence of shown in Fig. 2C (SEQ ID NO 7) and the Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 2D.

[0059] Figure 3A illustrates the expression levels in Parkinson's and Alzheimer's brain tissue and white blood cells of a leukemia patient of the GPCR gi6671985_GS_nt9.2 having the polynucleotide sequence shown in Fig. 3B (SEQ ID NO 3), the amino acid sequence of shown in Fig. 3C (SEQ ID NO 8) and the Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 3D.

[0060] Figure 4A illustrates the expression levels in Parkinson's and Alzheimer's brain tissue and white blood cells of a leukemia patient of the GPCR gi5791525_GS_nt10 having the polynucleotide sequence shown in Fig. 4B (SEQ ID NO 4), the amino acid sequence of shown in Fig. 4C (SEQ ID NO 9) and the Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 4D.

[0061] Figure 5A illustrates the expression levels in Parkinson's and Alzheimer's brain tissue and white blood cells of a leukemia patient of the GPCR gi5791525_GS_nt8 having the polynucleotide sequence shown in Fig. 5B (SEQ ID NO 5), the amino acid sequence of shown in Fig. 5C (SEQ ID NO 10) and the Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 5D.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[0062] The present invention provides novel methods and compositions involving GPCR polypeptides and polynucleotides specifically over-expressed or under-expressed in diseased tissues, specifically (and by example) Parkinson's disease, Alzheimer's disease and leukemia patients. The invention can be used in various aspects of genome analysis that finds utility in both basic biological research and medical diagnosis and therapeutics.

1. DEFINITIONS

[0063] A "polynucleotide" is a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example proteins (including for e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.),those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide.

[0064] The scope of the invention with respect to polynucleotide compositions useful in the methods described herein includes, but is not necessarily limited to, polynucleotides having a sequence set forth in any one of the polynucleotide sequences Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) provided herein or other biological sources (particularly human sources) or by hybridization to the above mentioned sequences under stringent conditions (particularly conditions of high stringency); genes corresponding to the provided polynucleotides; variants of the provided polynucleotides and their corresponding genes particularly those variants that retain a biological activity of the encoded gene product (*e.g.*, a biological activity ascribed to a gene product corresponding to the provided polynucleotides as a result of the assignment of the gene product to a protein family(ies) and/or identification of a functional domain present in the gene product). Other polynucleotide compositions contemplated by and within the scope of the present invention will be readily apparent to one of ordinary skill in the art when provided with the disclosure here.

[0065] "Polynucleotide" and "nucleic acid" are used herein inter-changably with reference to nucleic acids of the composition and is not intended to be limiting as to the

length or structure of the nucleic acid unless specifically indicted. As used herein, "polynucleotide" means a polymeric form of nucleotides of any length equal to or longer than a dimer of nucleotides, which contain deoxyribonucleotides, ribonucleotides, and/or their analogs. The terms "polynucleotide", "oligonucleotide" and "nucleotide" as used herein are used interchangeably. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The term "polynucleotide" includes double- or single-stranded molecules. Unless otherwise specified or required, any embodiment of the invention described herein that includes a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double stranded form. The following are non-limiting examples of polynucleotides: a gene or gene fragment, exons, introns, mRNA, tRNA, rRNA, ribozymes, codon, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, any isolated DNA from any source, any isolated RNA from any sequence, nucleic acid probes, and primers.

[0066] A polynucleotide or polynucleotide region has a certain percentage (for example, 80%, 85%, 90%, or 95%) of "sequence identity" to another sequence means that, when aligned, that percentage of bases are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in *Current Protocols in Molecular Biology* (F.M. Ausubel et al., eds., 1987) Supplement 30, section 7.7.18. A preferred alignment program is GCG Gap (Genetics Computer Group, Wisconsin, Suite Version 10.1), preferably using default parameters, such as: open gap = 3; extend gap = 1.

[0067] Polynucleotides contemplated by the invention also include naturally occurring variants of the nucleotide sequences (*e.g.*, degenerate variants, allelic variants, *etc.*). Variants of the polynucleotides contemplated by the invention are identified by hybridization of putative variants with the polynucleotide sequences disclosed herein, preferably by hybridization under stringent conditions. For example, by using appropriate wash conditions, variants of the polynucleotides described herein can be identified where the allelic variant exhibits at most about 25-30% base pair (bp) mismatches relative to the selected polynucleotide probe. In general, allelic variants contain 15-25% bp mismatches,

and can contain as little as even 5-15%, or 2-5%, or 1-2% bp mismatches, as well as a single bp mismatch.

[0068] The invention also encompasses homologues corresponding to any one of the polynucleotide sequences provided herein, where the source of homologous genes can be any mammalian species, *e.g.*, primate species, particularly human; rodents, such as rats, etc. Between mammalian species, *e.g.*, human and primate, homologues generally have substantial sequence similarity, *e.g.*, at least 75% sequence identity, usually at least 90%, more usually at least 95% between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, domain, *etc.* A reference sequence will usually be at least about 18 contiguous nucleotides long, more usually at least about 30 nucleotides long, and may extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art.

[0069] The subject nucleic acids can be cDNAs or genomic DNAs, as well as fragments thereof, particularly fragments that encode a biologically active gene product and/or are useful in the methods disclosed herein (*e.g.*, in diagnosis, as a unique identifier of a differentially expressed gene of interest, *etc.*). The term "codon" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons and 3' and 5' non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns, when present, being removed by nuclear RNA splicing, to create a continuous open reading frame encoding a polypeptide. mRNA species can also exist with both exons and introns, where the introns may be removed by alternative splicing. Furthermore it should be noted that different species of mRNAs encoded by the same genomic sequence can exist at varying levels in a cell, and detection of these various levels of mRNA species can be indicative of differential expression of the encoded gene product in the cell.

[0070] A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the

introns that are normally present in a native chromosome. It can further include the 3' and 5' untranslated regions found in the mature mRNA. It can further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, *etc.*, including about 1 kb, but possibly more, of flanking genomic DNA at either the 5' and 3' end of the transcribed region. The genomic DNA can be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' and 5', or internal regulatory sequences as sometimes found in introns, contains sequences required for proper tissue, stage-specific, or disease-state specific expression.

[0071] As used herein, the term "vector" refers to a polynucleotide construct designed for transduction/transfection of one or more cell types. Vectors may be, for example, "cloning vectors" which are designed for isolation, propagation and replication of inserted nucleotides, "expression vectors" which are designed for expression of a nucleotide sequence in a host cell, or a "viral vector" which is designed to result in the production of a recombinant virus or virus-like particle, or "shuttle vectors", which comprise the attributes of more than one type of vector.

[0072] A "host cell" includes an individual cell or cell culture which can be or has been a recipient of exogenous polynucleotides. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or change. A host cell includes cells transfected or infected *in vivo* or *in vitro* with a polynucleotide of this invention.

[0073] "Expression" includes transcription and/or translation.

[0074] The present invention further provides polypeptides, such as those comprising sequences shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10) and those encoded by polynucleotides shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) that are differentially expressed in tissue samples. The polypeptides can be used to generate antibodies specific for a polypeptide associated with Parkinson's disease, Alzheimer's disease and leukemia, which antibodies are in turn useful in diagnostic

methods, prognostic methods, therametric methods, and the like as discussed in more detail herein. Polypeptides are also useful as targets for therapeutic intervention, as discussed in more detail herein. The invention includes within its scope a polypeptide encoded by a polynucleotide having the sequence of any one of the polynucleotide sequences provided herein, or a variant thereof. The polypeptides of the present invention encompass polypeptides made by any method including isolation from a tissue or cell line source, recombinant expression, or chemical synthesis.

[0075] In general, the term "polypeptide" as used herein refers to both the full length polypeptide encoded by the recited polynucleotide, the polypeptide encoded by the gene represented by the recited polynucleotide, as well as portions or fragments thereof. The present invention encompasses variants of the naturally occurring proteins, wherein such variants are homologous or substantially similar to the naturally occurring protein, and can be of an origin of the same or different species as the naturally occurring protein (*e.g.*, human, murine, or some other species that naturally expresses the recited polypeptide, usually a mammalian species). In general, variant polypeptides have a sequence that has at least about 80%, usually at least about 90%, and more usually at least about 99% sequence identity with a differentially expressed polypeptide described herein, as determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is taught in Smith and Waterman, *Adv. Appl. Math.* (1981) 2: 482-489.. The variant polypeptides can be naturally or non-naturally glycosylated, *i.e.*, the polypeptide has a glycosylation pattern that differs from the glycosylation pattern found in the corresponding naturally occurring protein.

[0076] The invention also encompasses homologues of the disclosed polypeptides (or fragments thereof) where the homologues are isolated from other mammalian species, *e.g.* rodents, such as mice, rats; and primates and humans. By "homologue" is meant a polypeptide having at least about 35%, usually at least about 40% and more usually at least about 60% amino acid sequence identity to a particular differentially expressed protein as identified above, where sequence identity is determined by the Smith-

Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is taught in Smith and Waterman, *Adv. Appl. Math.* (1981) 2: 482-489. In general, the polypeptides of the subject invention are provided in a non-naturally occurring environment, *e.g.* are separated from their naturally occurring environment. In certain embodiments, the subject protein is present in a composition that is enriched for the protein as compared to a control. As such, purified polypeptide is provided, where by purified is meant that the protein is present in a composition that is substantially free of other expressed polypeptides, where by substantially free is meant that less than 90%, usually less than 60% and more usually less than 50% of the composition is made up of other expressed polypeptides.

[0077] Also within the scope of the invention are variants. Variants of polypeptides include mutants, fragments, and fusions. Mutants can include amino acid substitutions, additions or deletions. The amino acid substitutions can be conservative amino acid substitutions or substitutions to eliminate non-essential amino acids, such as to alter a glycosylation site, a phosphorylation site or an acetylation site, or to minimize misfolding by substitution or deletion of one or more cysteine residues that are not necessary for function. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/ hydrophilicity, and/or steric bulk of the amino acid substituted. Variants can be designed so as to retain or have enhanced biological activity of a particular region of the protein (*e.g.*, a functional domain and/or, where the polypeptide is a member of a protein family, a region associated with a consensus sequence). Selection of amino acid alterations for production of variants can be based upon the accessibility (interior vs. exterior) of the amino acid (see, e.g., Go *et al*, *Int. J. Peptide Protein Res.* (1980) *15*:211), the thermostability of the variant polypeptide (see, e.g., Querol *et al.*, *Prot. Eng.* (1996) *9*:265), desired glycosylation sites (see, e.g., Olsen and Thomsen, *J. Gen. Microbiol.* (1991) *137*:579), desired disulfide bridges (see, e.g., Clarke *et al.*, *Biochemistry* (1993) *32*:4322; and Wakarchuk *et al.*, *Protein Eng.* (1994) 7:1379), desired metal binding sites (see, e.g., Toma *et al.*, *Biochemistry* (1991) *30*:97, and Haezerbrouck *et al.*, *Protein Eng.* (1993) *6*:643), and desired substitutions with in proline

loops (see, e.g., Masul *et al.*, *Appl. Env. Microbiol.* (1994) *60*:3579). Cysteine-depleted muteins can be produced as disclosed in U.S. Pat. No. 4,959,314.

[0078] Variants also include fragments of the polypeptides disclosed herein, particularly biologically active fragments and/or fragments corresponding to functional domains. Fragments of interest will typically be at least about 8 amino acids (aa) 10 aa, 15 aa, 20 aa, 25 aa, 30 aa, 35 aa, 40 aa, to at least about 45 aa in length, usually at least about 50 aa in length, at least about 75 aa, at least about 100 aa, at least about 125 aa, at least about 150 aa in length, at least about 200 aa, at least about 300 aa, at least about 400 aa and can be as long as 500 aa in length or longer, but will usually not exceed about 1000 aa in length, where the fragment will have a stretch of amino acids that is identical to a polypeptide encoded by a polynucleotide having a sequence of any one of the polynucleotide sequences provided herein, or a homologue thereof. The protein variants described herein are encoded by polynucleotides that are within the scope of the invention. The genetic code can be used to select the appropriate codons to construct the corresponding variants.

[0079] While the over-expression of the polynucleotides associated with Parkinson's disease, Alzheimer's disease and leukemia is observed, elevated levels of expression of the polypeptides encoded by these polynucleotides may likely play a role in Parkinson's disease, Alzheimer's disease and leukemia.

[0080] A "G-protein" or "GTP-binding protein" is one of a superfamily of proteins that function in, for example, signal transduction (e.g. the G-protein associated with the b-adrenergic receptor), polymerization (e.g. tubulin), ribosomal protein synthesis (e.g. the translocase), cell differentiation (ras proteins) and intracellular transport of proteins, vesicles or cytoskeletal elements (e.g. dynamin). The common functional feature of the family is an affinity for a target protein in the presence of GTP which is lost upon hydrolysis to GDP. (*see*, Bourne, H.R., Sanders, D.A. and McCormick, F. (1990) Nature (London) **348**, 125-132; Hilgenfeld, R. (1995) Curr. Opin. Struct. Biol. **5**, 810-817).

[0081] The class of receptor termed "G-protein coupled receptors (GPCRs)" couple signal transduction to G-proteins. GPCRs are coupled, inside the cell, to GTP-binding

and hydrolyzing proteins (termed G-proteins). Receptors of the class that interact with G-proteins all have a structure that is characterized by 7 transmembrane (7-TM) spanning domains. These receptors are also termed *serpentine* receptors. Examples of this class are the adrenergic receptors, odorant receptors, and certain hormone receptors (e.g. glucagon, angiotensin, vasopressin and bradykinin).

[0082] A "microarray" is a linear or two-dimensional array of preferably discrete regions, each having a defined area, formed on the surface of a solid support. The density of the discrete regions on a microarray is determined by the total numbers of target polynucleotides to be detected on the surface of a single solid phase support, preferably at least about 50/cm^2, more preferably at least about 100/cm^2, even more preferably at least about 500/cm^2, and still more preferably at least about 1,000/cm^2. As used herein, a DNA microarray is an array of oligonucleotide primers placed on a chip or other surfaces used to amplify or clone target polynucleotides. Since the position of each particular group of primers in the array is known, the identities of the target polynucleotides can be determined based on their binding to a particular position in the microarray.

[0083] The term "label" refers to a composition capable of producing a detectable signal indicative of the presence of the target polynucleotide in an assay sample. Suitable labels include radioisotopes, nucleotide chromophores, enzymes, substrates, fluorescent molecules, chemiluminescent moieties, magnetic particles, bioluminescent moieties, and the like. As such, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means.

[0084] The term "support" refers to conventional supports such as beads, particles, dipsticks, fibers, filters, membranes and silane or silicate supports such as glass slides.

[0085] As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, cells (including but not limited to blood cells), tumors, organs, and also samples of in vitro cell culture constituents.

[0086] The term "biological sources" as used herein refers to the sources from which the target polynucleotides are derived from. The source can be of any form of "sample" as described above, including but not limited to, cell, tissue or fluid. "Different biological sources" can refer to different cells/tissues/organs of the same individual, or cells/tissues/organs from different individuals of the same species, or cells/tissues/organs from different species.

2. METHODS AND MATERIALS

Selecting oligonucleotide probes

[0087] The invention provides a prepared solid support comprising immobilized and separate groups of oligonucleotide probes. Each probe group corresponds to a particular region within the reference sequence, and contains at least four sets of probes with the first set being exactly complementary to the particular region of the reference sequence, and the other three sets being identical to the first set but for the most 3'-end nucleotide. For example, for an A nucleotide in the reference sequence, the corresponding probe from the first set has a T at its most 3'-end, while the additional three probe sets have at their most 3'-end an A, C, or G, a different nucleotide in each set. The length of the four probe sets are preferred, although not necessary, to be the same. The probes can be selected or designed using for example a standard PCR probe selection program such as Probe3 from Massachusetts Institute of Technology (MIT).

[0088] The solid phase support can provide an areas of about 5 to about 100 square micrometers, on which up to about 100,000 groups of probes can be immobilized in discrete areas according to a predetermined pattern. The prepared solid support can have an associated written or electronic record of the sequence of the probe or probe pairs at any given location on the support, and thus the location on the support of an amplified target can be identified as well.

[0089] The number of probes within each group corresponding to a particular region of the reference sequence can be determined and limited by the needs of the subsequent planned amplification reaction on the microarray. Thus, for example, the number of

probes deemed necessary for conducting an PCR at a specific site on the microarray, given especially the reaction volume and expected number of target template polynucleotide molecules, and the proposed number of cycles of PCR, will help determine exactly how much oligonucleotide probe copies to apply as a group at each location on the support to ensure successful reactions. Preferably, the amounts of probes (i.e. probe molecule numbers or probe concentration) will be about the same at each provided location on a given solid support (e.g. in a DNA microarray format having from 1000, to 10,000, up to about 100,000 groups of probes to amplify or detect up to about 100,000 regions of the target polynucleotide).

[0090] The solid support can be prepared with probe sequences for a particular application based on the polynucleotides to be detected. The oligonucleotide probes can be of any length suitable for a particular PCR, especially considering the sequence and quality of the target polynucleotides to be amplified. As an example, the probes can be from about 4 to about 30 nucleotides in length.

[0091] It is understood that a nucleic acid probe of the present invention may contain minor deletions, additions and/or substitutions of nucleic acid bases, to the extent that such alterations do not negatively affect the yield or product obtained to a significant degree.

[0092] Oligonucleotide probes can include the naturally-occurring heterocyclic bases normally found in nucleic acids (uracil, cytosine, thymine, adenine and guanine), as well as modified bases and base analogues. Any modified base or base analogue compatible with hybridization of the probe to a target sequence is useful in the practice of the invention.

[0093] The sugar or glycoside portion of the probe can comprise deoxyribose, ribose, and/or modified forms of these sugars, such as, for example, 2'-O-alkyl ribose. In a preferred embodiment, the sugar moiety is 2'-deoxyribose; however, any sugar moiety that is compatible with the ability of the probe to hybridize to a target sequence can be used.

[0094] In one embodiment, the nucleoside units of the probe are linked by a phosphodiester backbone, as is well known in the art. In additional embodiments, internucleotide linkages can include any linkage known to one of skill in the art that is compatible with specific hybridization of the probe including, but not limited to phosphorothioate, methylphosphonate, sulfamate (*e.g.*, U.S. Patent No. 5,470,967) and polyamide (*i.e.*, peptide nucleic acids). Peptide nucleic acids are described in Nielsen *et al.* (1991) *Science* 254: 1497-1500, U.S. Patent No. 5,714,331, and Nielsen (1999) *Curr. Opin. Biotechnol.* 10:71-75.

[0095] In certain embodiments, the probe can be a chimeric molecule; *i.e.*, can comprise more than one type of base or sugar subunit, and/or the linkages can be of more than one type within the same probe.

[0096] The probe can comprise a moiety to facilitate hybridization to its target sequence, as are known in the art, for example, intercalators and/or minor groove binders.

[0097] Variations of the bases, sugars, and internucleoside backbone, as well as the presence of any pendant group on the probe, will be compatible with the ability of the probe to bind, in a sequence-specific fashion, with its target sequence. A large number of structural modifications, both known and to be developed, are possible within these bounds. Moreover, synthetic methods for preparing the various heterocyclic bases, sugars, nucleosides and nucleotides which form the probe, and preparation of oligonucleotides of specific predetermined sequence, are well-developed and known in the art. A preferred method for oligonucleotide synthesis incorporates the teaching of U.S. Patent No. 5,419,966.

[0098] The oligonucleotide probes can be designed with any special additional moieties or sequences that will aid and facilitate a particular PCR or subsequent manipulations, e.g. isolation of the amplified target polynucleotides. For example, a probe can comprise sequences in addition to those that are complementary to the target sequence. Such sequences are normally upstream (*i.e.*, to the 5'-side) of the target-complementary sequences in the probe. For example, sequences comprising one or more restriction enzyme recognition sites (so-called "linkers" or "adapters"), when present in a probe

upstream of target-complementary sequences, facilitate cloning and subsequent manipulation of an amplification product. Other useful sequences for inclusion in a probe include those complementary to a sequencing probe and those specifying a promoter for a bacteriophage RNA polymerase, such as, for example, T3 RNA polymerase, T7 RNA polymerase and/or SP6 RNA polymerase.

[0099] In one aspect of the invention, the microarray probes are defined by a tiling method to cover an entire region of interest in the target polynucleotide. For example, a first group of probes are designed so that the sequence of each probe therein corresponds to the most 5'-portion of the region of interest; a second group of probes have sequence that is "shifted" from the first group by one nucleotide towards the 3'-end of the region; and a third group of probes have sequence that is "shifted" from the second group by one nucleotide toward the 3'-end of the region, and etc. In theory, then, the number of groups of probes equals to the number of nucleotides in the region of interest. Of course, within each group of probes that correspond to a particular portion of the region, there are at least four sets of probes with four different 3'-ends as described above. When multiple target polynucleotides are to be detected according to the present invention, each probe group corresponding to a particular target polynucleotide is resided in a discrete area of the microarray.

Solid Phase Supports

[0100] 5'-modified oligonucleotides representing novel GPCR sequences and optimized for DNA microarray analysis are spotted on DNA microarray slides. The microarray slides of the present invention can be of any solid materials and structures suitable for supporting nucleotide hybridization and synthesis. Preferably, the solid phase support comprises at least one substantially rigid surface on which the primers can be immobilized and the PCR reaction performed. The solid phase support can be made of, for example, glass, synthetic polymer, plastic, hard non-mesh nylon or ceramic. Other suitable solid support materials are known and readily available to those of skill in the art. The size of the solid support can be any of the standard microarray sizes, useful for DNA microarray technology, and the size may be tailored to fit the particular machine being

used to conduct a reaction of the invention. Methods and materials for derivatization of solid phase supports for the purpose of immobilizing oligonucleotides are known to those skill in the art and described in, for example, U.S. Pat. No. 5,919,523, the disclosure of which is incorporated herein by reference.

Probe Immobilization

[0101] The oligonucleotide primers of the invention are affixed, immobilized, provided, and/or applied to the surface of the solid support using any available means to fix, immobilize, provide and/or apply the oligonucleotides at a particular location on the solid support. For example, photolithography (Affymetrix, Santa Clara, CA) can be used to apply the oligonucleotide primers at particular position on a chip or solid support, as described in the U.S. patents, USPN 5,919,523, USPN 5,837,832, USPN 5,831,070, and USPN 5,770,722, which are incorporated herein by reference. The oligonucleotide primers may also be applied to a solid support as described in Brown and Shalon, USPN 5,807,522 (1998). Additionally, the primers may be applied to a solid support using a robotic system, such as one manufactured by Genetic MicroSystems (Woburn, MA), GeneMachines (San Carlos, CA) or Cartesian Technologies (Irvine, CA).

PCR Amplifications

[0102] In practicing the invention, a reaction mixture comprising the appropriate target polynucleotides mixed with the reagents necessary for conducting the polymerase chain reaction (PCR) are placed in contact with each immobilized primer pair or single primer population on the solid support. The appropriate target polynucleotides can be double stranded DNA, single stranded codon generated by reverse transcription of RNA templates, or mRNA population. The reaction mixture contains an enzyme for facilitating the synthesis of a polynucleotide strand complementary to a target strand. Suitable polymerases include thermostable polymerase enzymes, such as Taq DNA polymerase, TthI DNA polymerase, Tne DNA polymerase, Tma DNA polymerase, Pfu DNA polymerase, Vent DNA polymerase or any other thermostable DNA polymerase. The reaction mixture can also contain a label molecule capable of being incorporated into the

nascent strands during polymerase chain reaction so that the amplified products can be detected on the solid support after the PCR. The label can be detected directly or indirectly according to methods well known in the art. Suitable labels for direct detection can be any fluorescent molecules such as fluorescein isothiocyanate, Texas red or rhodamine. Molecules facilitating indirect detection, such as biotin or digoxigenin, can also be incorporated into the nascent strands during the PCR. Biotin can be subsequently detected by binding to a labeled streptavidin or a labeled anti-biotin antibody. Likewise, incorporated digoxigenin can be detected by a labeled or unlabeled anti-digoxigenin antibody, and the unlabeled anti-digoxigenin antibody can be detected by binding a labeled anti-anti-digoxigenin antibody.

Labeling and Detection

[0103] Detecting the amplified or labeled target polynucleotides can be conducted by standard methods used to detect the labeled sequences, including for example, detecting labels that have been incorporated into the amplified or newly synthesized DNA strands. Thus, for example fluorescent labels or radiolabels can be detected directly. Other labeling techniques may require that a label such as biotin or digoxigenin that is incorporated into the DNA during strand synthesis be detected by an antibody or other binding molecule (e.g. streptavidin) that is either labeled or which can bind a labeled molecule itself, for example, a labeled molecule can be e.g. an anti-streptavidin antibody or anti-digoxigenin antibody conjugated to either a fluorescent molecule (e.g. fluorescein isothiocyanate, Texas red and rhodamine), or conjugated to an enzymatically activatable molecule. Whatever the label on the newly synthesized molecules, and whether the label is directly in the DNA or conjugated to a molecule that binds the DNA (or binds a molecule that binds the DNA), the labels (e.g. fluorescent, enzymatic, chemiluminescent, or colorimetric) can be detected by a laser scanner or a CCD camera, or X-ray film, depending on the label, or other appropriate means for detecting a particular label.

[0104] The target polynucleotides can be detected by using labeled nucleotides (e.g. dNTP-fluorescent label for direct labeling; dNTP-biotin or dNTP-digoxigenin for indirect labeling) are incorporated into amplified DNA during the PCR. For indirectly labeled

DNA, the detection is carried out by fluorescence or other enzyme conjugated streptavidin or anti-digoxigenin antibodies. The PCR method employs detection of the polynucleotides by detecting incorporated label in the newly synthesized complements to the polynucleotide targets. For this purpose, any label that can be incorporated into DNA as it is synthesized can be used, e.g. fluoro-dNTP, biotin-dNTP, or digoxigenin-dNTP, as described above and are known in the art. PCR amplification conducted using one or more universal primers in solution provides the option to detect the amplified targets at locations on the solid support by detecting the universal primers. Thus, where more than one universal primer is used, target strands from different sources can be differentially detected on the solid support.

[0105] Examples of suitable fluorescent labels include fluorescein (FITC), 5,6-carboxymethyl fluorescein, Texas red, nitrobenz-2-oxa-1,3-diazol-4-yl (NBD), coumarin, dansyl chloride, rhodamine, 4'-6-diamidino-2-phenylinodole (DAPI), and the cyanine dyes Cy3, Cy3.5, Cy5, Cy5.5 and Cy7. Preferred fluorescent labels are fluorescein (5-carboxyfluorescein-N-hydroxysuccinimide ester) and rhodamine (5,6-tetramethyl rhodamine). Preferred fluorescent labels for combinatorial multicolor coding are FITC and the cyanine dyes Cy3, Cy3.5, Cy5, Cy5.5 and Cy7. The absorption and emission maxima, respectively, for these fluors are: FITC (490 nm; 520 nm), Cy3 (554 nm; 568 nm), Cy3.5 (581 nm; 588 nm), Cy5 (652 nm: 672 nm), Cy5.5 (682 nm; 703 nm) and Cy7 (755 nm; 778 nm), thus allowing their simultaneous detection. The fluorescent labels can be obtained from a variety of commercial sources, including Molecular Probes, Eugene, OR and Research Organics, Cleveland, Ohio.

[0106] Labeled nucleotides are preferred form of detection label since they can be directly incorporated during synthesis. Examples of detection labels that can be incorporated into amplified DNA or RNA include nucleotide analogs such as BrdUrd (Hoy and Schimke, Mutation Research 290:217-230 (1993)), BrUTP (Wansick et al., J. Cell Biology 122:283-293 (1993)) and nucleotides modified with biotin (Langer et al., Proc. Natl. Acad. Sci. USA 78:6633 (1981)) or with suitable haptens such as digoxigenin (Kerkhof, Anal. Biochem. 205:359-364 (1992)). Suitable fluorescence-labeled nucleotides are Fluorescein-isothiocyanate-dUTP, Cyanine-3-dUTP and Cyanine-5-dUTP

(Yu et al., Nucleic Acids Res., 22:3226-3232 (1994)). A preferred nucleotide analog detection label for DNA is BrdUrd (BUDR triphosphate, Sigma), and a preferred nucleotide analog detection label for RNA is Biotin-16-uridine-5'-triphosphate (Biotin-16-dUTP, Boehringer Mannheim). Fluorescein, Cy3, and Cy5 can be linked to dUTP for direct labeling. Cy3.5 and Cy7 are available as avidin or anti-digoxigenin conjugates for secondary detection of biotin- or digoxigenin-labeled probes.

Comparing Differential Expression of Target Polynucleotides from Different Biological Sources

[0107] For most studies involving gene expression, RNA is isolated from specific tissue samples. This RNA is then subjected to reverse transcription using oligo-dT primers and fluorescently labeled dNTPs (usually Cy3 or Cy5 labeled dCTP) resulting in a DNA probe that is fluorescently labeled and has a complementary sequence to the original mRNA. The next step is to hybridize the probe to the immobilized target DNA attached to the microarray. This is done by denaturing the probe with heat or a mild base to reduce secondary structures that may have formed and applying it onto the microarray. A cover slip is applied to the array to ensure even distribution of the probe. The array is placed in a warm, humidified chamber overnight to allow the single stranded probe DNA to bind to its complementary single stranded target. The microarray is then removed and washed to remove any nonspecifically bound probe. The arrays are then imaged with a confocal laser scanner. The scanner contains 2 lasers tuned to excite the dye incorporated into the DNA probe and a corresponding filter set to select out excitation emission from the dye (Cy3 or Cy5). The ability to image two fluorescent signals allows for two different RNA samples to be hybridized and directly compared on the same array. This excitation emission signal is recorded via a photomultiplier tube (PMT), digitized, and sent to the computer for later analysis. By examining the intensity of a spot's fluorescence, and the ratio of fluorescence between spots, it possible to determine whether a specific gene is being expressed and the relative expression level of the gene between samples. Other available means for labeling and detecting probes, such as with radioisotopes, enzymes,

antibodies, biotin, avidin and like materials known in the art, are within the contemplated means of executing the process.

[0108] Solid phase amplification methods can be used to detect and compare gene expressions in different biological sources. The immobilized primers are used in combination with solution phase primers to conduct amplification reactions. Different sources can be different tissues or cells of the same subject. Alternatively, different sources can be comparable tissues of two or more different subjects of same species, e.g., one from a healthy control, and another from a patient. In another embodiment, the different sources are two or more different species or different animals, such as one of human and another of mouse.

[0109] As an illustration, the original probe polynucleotides from different biological sources are total mRNAs expressed therein. Methods and materials that are known in the art are used to isolate total mRNAs from each source. The total pool of isolated mRNA from each biological source is then used to prepare a batch of specifically tagged codon that are to be used as "probe polynucleotides" for subsequent amplification. Each batch of the reverse-transcribed cDNAs are tagged with a specific sequence tag at their 3-ends. The specific sequence tag is not present in any of the unmodified probe polynucleotides. For example, if the probe polynucleotides to be detected are from human cells/tissues, the sequence tag can be derived from a bacteria or viral genome such that the sequence does not anneal under hybridization/amplification conditions with the human sequences that are transcribed into mRNA. In this way, the sequence tags from one batch will not cause artifactual amplification of another batch due to cross-hybridization.

[0110] Furthermore, the sequence tag for each batch of codon probes is different from that for another batch of codon probes so that they can be compared. The sequence tag can be introduced into the reverse-transcribed codon by using a specially designed primer for reverse transcription. For example, a primer can have a poly-dT portion at its 3'-end and a bacterial SP6 sequence at its 5'-end. During reverse transcription, mRNAs serve as template for a codon synthesis initiated from the 5'-end of the poly-dT portion that anneals with the poly-A tail of the mRNA templates. The resulting codon products then

have at their 5'-end a SP6 "sequence tag," which is unique to this batch of cDNAs. Similarly, a different batch of cDNAs from another source can be "tagged" with a different sequence tag, such as a bacterial T7 sequence.

[0111] Two batches of cDNAs differentially tagged with, for example, SP6 and T7 are mixed together for amplification and detection. Present in the amplification reaction mixture are differentially labeled, free SP6 and T7 sequence tags. For example, the two sequence tags can be labeled either with two different fluorescent dyes (e.g., one red dye and one green dye) for direct detection or, alternatively, with two chemical moieties (e.g., one biotin and one digoxigenin) for subsequent color detection. It is important that the labels do not occur at the 3'-end of the sequence tags so that the sequence tags can latter serve as primers in amplification reaction.

[0112] A preferred amplification means for the invention is PCR reaction. The mixture of two differentially tagged batches of cDNAs are contacted with an array of multiple groups of specific primer targets, with each group corresponding to a particular probe codon as described above. In the initial round of PCR, the immobilized primers anneal with probe polynucleotides from both sources and synthesize a nascent complementary strand under conditions sufficient for chain elongation. The nascent complementary strand spans through the probe sequence region and contains at its 3'-end a sequence complementary to the sequence tag at the end of the probe codon template. For example, a first nascent strand has a 3'-end complementary to the SP6 sequence if it was amplified on a SP6-tagged codon template from source 1; and a second nascent strand has a 3'-end complementary to the T7 sequence if it was amplified on a T7-tagged codon template from source 2. Thus, each nascent strand immobilized on the solid phase array "inherits" the specific sequence tag specific to its source.

[0113] In the subsequent rounds of PCR, the first set of nascent strands tagged for different sources serve as templates for the synthesis of a second set of nascent strands. This time, the initial primers of the PCR are the differentially labeled free sequence tags in solution, such as fluorescein-labeled SP6 primers and lissamine-labeled T7 primers. Accordingly, the second set of nascent strands extended from the labeled sequence tags

are differentially labeled corresponding to the original sources of the probe codon templates in the original round of PCR reaction. After washing off the unbound reagents and original templates without denaturation, each immobilized primer site will have a double stranded polynucleotide having on one strand a label indicating the original biological source. As such, the final detection of different labels will reveal the presence and abundance of particular probe polynucleotides in different biological sources.

3. EXAMPLES

[0114] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all and only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Celsius, and pressure is at or near atmospheric.

Example 1. Generation of oligonucleotide probe arrays

[0115] Putative GPCR sequences, based on well-known GPCR characteristics including the 7 trans-membrane (7-TM) structure, the "DRY" (amino acids Asp-Arg-Tyr) motif and homologies to known human GPCR sequences were searched among human genomic sequences in public database GenBank (NCBI). The GENSCAN bioinformatics software (Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268, 78-94. (1997); v.1999 version under license from Stanford University, CA) and in house filtering software were used for the search.

[0116] A semi-automatic protein families database of alignments and hidden Markov models (PFAM) software (Pfam, release 5) was used to analyze the putative GPCR sequences and capture "all human GPCR sequences" from the input. (*see generally,* Sonnhammer et al. (1997) Protein 28:405-420).

[0117] All "known" human GPCR sequences in GenBank were then subtracted from the "all human GPCR" sequences obtained from the Pfam analysis and the remaining sequences were designated as "novel" GPCR sequences. Amino acid sequences corresponding to the "novel" GPCR sequences were then analyzed by a Kyte-Doolittle hydropathicity plot. (Kyte, J. and Doolittle, R. F. "A Simple Method for Displaying the Hydropathic Character of a Protein", J. Mol. Biol. 157:105-132 (1982)). Kyte-Doolittle hydrophobicity analysis identified the seven hydrophobic domains characteristic of G-protein coupled receptors and "DRY" motif was searched for by scanning the amino acid sequence.

[0118] Novel sequences that matched the sequence and structural criteria for GPCR were targeted for full-length clone isolation from genomic and codon libraries. The full-length clones were sequenced, and Kyte-Doolittle hydropathicity plots and "DRY" motif amino acid sequence analyses were subsequently performed to confirm their status in the GPCR family. The amino acid sequence of the GPCR polynucleotide gi6863021_GS_nt7 (Fig. 1B; SEQ ID NO 1) is shown in Fig. 1C (SEQ ID NO 6) and the corresponding Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 1D. The amino acid sequence of the GPCR polynucleotide gi6453999_GS_nt6 (Fig. 2B; SEQ ID NO 2) is shown in Fig. 2C (SEQ ID NO 7) and the corresponding Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 2D. The amino acid sequence of the GPCR polynucleotide gi6671985_GS_nt9.2 (Fig. 3B; SEQ ID NO 3) is shown in Fig. 3C (SEQ ID NO 8) and the corresponding Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 3D. The amino acid sequence of the GPCR polynucleotide gi5791525_GS_nt10 (Fig. 4B; SEQ ID NO 4) is shown in Fig. 4C (SEQ ID NO 9) and the corresponding Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 4D. The amino acid sequence of the GPCR polynucleotide gi5791525_GS_nt8 (Fig. 5B; SEQ ID NO 5) is shown in Fig. 5C (SEQ ID NO 10) and the corresponding Kyte-Doolittle hydropathicity plots and "DRY" motif is shown in Fig. 5D.

[0119] Summary of polynucleotides described herein: Table 1 provides a summary of polynucleotides isolated as described above and identified as corresponding to a differentially expressed gene (see Example 2 below). Specifically, Table 1 provides the

relative over-expression of a GPCR polynucleotide having a particular clone ID in Parkinson's disease, Alzheimer's disease and leukemia tissue samples.

Example 2: Detection of altered levels of GPCR RNA expressed in disease tissues using arrays.

[0120] GPCR-related sequences representing a variety of candidate genes to be screened for differential expression in Parkinson's disease, Alzheimer's disease and leukemia were assayed by hybridization on polynucleotide arrays. 5'-modified oligonucleotides representing the novel GPCR sequences and optimized for DNA microarray analysis were designed using a combination of GENSCAN, Pfam, Primer3 (release 0.9 (1998) MIT Whitehead Institute) and in-house informatics programs and were spotted onto reflective glass slides (Amersham) according to methods well known in the art . These microarrays were then used to study differential gene expression in selected cell lines and patient tissues. Normal tissues and tissues from Parkinson's disease, Alzheimer's disease and leukemia patients were processed to generate T7 RNA polymerase transcribed polynucleotides, which were, in turn, assessed for expression in the microarrays.

[0121] The microarrays were then used to study differential gene expression in various disease samples and appropriate normal control samples. Target polynucleotides were prepared from total RNA obtained by laser capture microdissection (LCM, Arcturus Engineering Inc., Mountain View, CA) of diseased tissue samples and normal tissue samples. Total RNA was first reverse transcribed into codon using a primer containing a T7 RNA polymerase promoter, followed by second strand DNA synthesis. The codon was then transcribed *in vitro* to produce antisense RNA using the T7 promoter-mediated expression (see, *e.g.*, Luo *et al.* (1999) *Nature Med* 5:117-122). In one embodiment, biotin labeled antisense RNA is generated from the RNA isolated from sample and control tissue by standard protocols and hybridized to the immobilized oligonucleotides representing the novel GPCRs.

[0122] Hybridization is detected by binding the biotinylated antibody with streptavidin which is in turn bound by a first anti-streptavidin antibody and the signals are amplified using a secondary antibody which binds the anti-streptavidin antibody and is conjugated

to the fluorescent label Cy3. The fluorescent label is detected by commercially available array scanners. These indirect immunofluorescence techniques are known in the art.

[0123] The array images generated by hybridization were captured by a laser array scanner (e.g., GenePix 4000B by Axon Instruments, Foster City, CA) and the data was analyzed using DNA array image analysis software (ImaGene 4.0, BioDiscovery, Inc.) and an in-house microarray data analysis software (Mergen).

[0124] In one embodiment, the procedure provides for fluorescent labeling of RNA. Probes are labeled by making fluorescently labeled codon from the RNA starting material. Fluorescently labeled cDNAs prepared from the diseased RNA sample are compared to fluorescently labeled cDNAs prepared from normal cell RNA sample. For example, in one embodiment the codon probes from the normal cells are labeled with Cy3 fluorescent dye (green) and the codon probes prepared from the diseased cells are labeled with Cy5 fluorescent dye (red).

[0125] The differential expression assay can be performed by mixing equal amounts of probes from disease cells and normal cells of the same patient. The arrays were prehybridized by incubation for about 2 hrs at 60°C in 5X SSC/0.2% SDS/1 mM EDTA, and then washed three times in water and twice in isopropanol. Following prehybridization of the array, the probe mixture was then hybridized to the array under conditions of high stringency (overnight at 42°C in 50% formamide, 5X SSC, and 0.2% SDS). After hybridization, the array was washed at 55°C three times as follows: (i) first wash in 1X SSC/0.2% SDS; (ii) second wash in 0.1X SSC/0.2% SDS; and (iii) third wash in 0.1X SSC.

[0126] The arrays were then scanned for green and/or red fluorescence using a Molecular Dynamics Generation III dual color laser-scanner/detector. The images were processed using the AutoGene Microarray Image Analysis System™ (BioDiscovery, Inc., Los Angeles, Calif.) and the data from each scan set was normalized. In one embodiment, the experiment was repeated, this time labeling the two probes with the opposite color in order to perform the assay in both "color directions." Each experiment was sometimes repeated with two more slides (one in each color direction). The data from each scan was

normalized, and the level fluorescence for each sequence on the array expressed as a ratio of the geometric mean of 8 replicate spots/genes from the four arrays or 4 replicate spots/gene from 2 arrays or some other permutation. Table 1 summarizes the results for gene products differentially expressed in the Parkinson's disease, Alzheimer's disease and leukemia tissue samples relative to normal cells.

[0127] As shown in Figures 1A, 2A, 3A, 4A and 5A and summarized in Table 1, four of the novel GPCR genes showed differential expression patterns in one or more tissue samples from Parkinson's disease, Alzheimer's disease and leukemia as compared to normal samples.

[0128] The GPCR polynucleotide gi6863021_GS_nt7 having the sequence shown in Fig. 1B (SEQ ID NO 1) is expressed at about a 7-fold reduced level in Alzheimer's tissue, and a 6-fold increased level in leukemia as shown in Fig. 1A.

[0129] The GPCR polynucleotide gi6453999_GS_nt6 having the sequence shown in Fig. 2B (SEQ ID NO 2) is expressed at about a 8-fold reduced level in Parkinson's tissue, a 5-fold reduced level in Alzheimer's tissue, and a 5-fold increased level in leukemia as shown in Fig. 2A.

[0130] The GPCR polynucleotide gi6671985_GS_nt9.2 having the sequence shown in Fig. 3B is expressed (SEQ ID NO 3) at about a 2-fold reduced level in Parkinson's tissue, a 4-fold reduced level in Alzheimer's tissue, and a 6-fold increased level in leukemia as shown in Fig. 3A.

[0131] The GPCR polynucleotide gi5791525_GS_nt10 having the sequence shown in Fig. 4B (SEQ ID NO 4) is expressed at about a 5-fold reduced level in Parkinson's tissue, a 6-fold reduced level in Alzheimer's tissue, and a 7-fold increased level in leukemia as shown in Fig. 4A.

[0132] The GPCR polynucleotide gi5791525_GS_nt8 having the sequence shown in Fig. 5B (SEQ ID NO 5) is expressed at about a 6-fold increased level in leukemia as shown in Fig. 5A

Table 1.

Clone ID No. (SEQ ID NO)	DRY motif at (aa residue/ total aa)	GPCR Expression in Diseased over Normal Tissue		
		Parkinson's Disease	Alzheimer's Disease	Leukemia
gi6863021_GS_nt7 (SEQ ID NO 1)	119/308	N/D	0.143 x normal	6 x normal
gi6453999_GS_nt6 (SEQ ID NO 2)	121/312	0.120 x normal	0.200 x normal	5 x normal
gi6671985_GS_nt9.2 (SEQ ID NO 3)	168/356	0.500 x normal	0.250 x normal	6 x normal
gi5791525_GS_nt10 (SEQ ID NO 4)	21/309	0.208 x normal	0.167 x normal	7 x normal
gi5791525_GS_nt8 (SEQ ID NO 5)	123/314	N/D	N/D	6 x normal

[0133] These genes are thus are possible targets for diagnosis and treatment of Parkinson's disease, Alzheimer's disease and leukemia and are also useful targets for drug development in the areas of CNS diseases.

Example 3: Expression of cloned polynucleotides in host cells.

[0134] To study the protein products of GPCR codon, restriction fragments from the GPCR codon are cloned into the expression vector pMT2 (Sambrook, et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press pp. 16.17-16.22 (1989)) and transfected into COS cells grown in DMEM supplemented with 10% FCS. Transfections are performed employing calcium phosphate techniques (Sambrook, et al (1989) pp. 16.32-16.40, supra) and cell lysates are prepared forty-eight hours after

transfection from both transfected and untransfected COS cells. Lysates are subjected to analysis by immunoblotting using anti-peptide antibody.

[0135] In immunoblotting experiments, preparation of cell lysates and electrophoresis are performed according to standard procedures. Protein concentration is determined using BioRad protein assay solutions. After semi-dry electrophoretic transfer to nitro-cellulose, the membranes are blocked in 500 mM NaCl, 20 mM Tris, pH 7.5, 0.05% Tween-20 (TTBS) with 5% dry milk. After washing in TTBS and incubation with secondary antibodies (Amersham), enhanced chemiluminescence (ECL) protocols (Amersham) are performed as described by the manufacturer to facilitate detection.

Example 4: Generation of antibodies against polypeptides.

[0136] Polypeptides, unique to GPCR are synthesized or isolated from bacterial or other (e.g., yeast, baculovirus) expression systems and conjugated to rabbit serum albumin (RSA) with m-maleimido benzoic acid N-hydroxysuccinimide ester (MBS) (Pierce, Rockford, Ill.). Immunization protocols with these peptides are performed according to standard methods. Initially, a pre-bleed of the rabbits is performed prior to immunization. The first immunization includes Freund's complete adjuvant and 500 μg conjugated peptide or 100 μg purified peptide. All subsequent immunizations, performed four weeks after the previous injection, include Freund's incomplete adjuvant with the same amount of protein. Bleeds are conducted seven to ten days after the immunizations.

[0137] For affinity purification of the antibodies, the corresponding GPCR polypeptide is conjugated to RSA with MBS, and coupled to CNBr-activated Sepharose (Pharmacia, Uppsala, Sweden). Antiserum is diluted 10-fold in 10 mM Tris-HCl, pH 7.5, and incubated overnight with the affinity matrix. After washing, bound antibodies are eluted from the resin with 100 mM glycine, pH 2.5.

Example 5: ELISA assay for Detecting GPCR-related sequences.

[0138] To test blood samples for antibodies that bind specifically to recombinantly produced GPCR antigens, the following procedure is employed. After the recombinant

GPCR-related proteins are purified, the recombinant protein is diluted in PBS to a concentration of 5 μg/ml (500 ng/100 μl). 100 microliters of the diluted antigen solution is added to each well of a 96-well Immulon 1 plate (Dynatech Laboratories, Chantilly, Va.), and the plate is then incubated for 1 hour at room temperature, or overnight at 4° C., and washed 3 times with 0.05% Tween 20 in PBS. Blocking to reduce nonspecific binding of antibodies is accomplished by adding to each well 200 μl of a 1% solution of bovine serum albumin in PBS/Tween 20 and incubation for 1 hour. After aspiration of the blocking solution, 100 μl of the primary antibody solution (anticoagulated whole blood, plasma, or serum), diluted in the range of 1/16 to 1/2048 in blocking solution, is added and incubated for 1 hour at room temperature or overnight at 4° C. The wells are then washed 3 times, and 100 μl of goat anti-human IgG antibody conjugated to horseradish peroxidase (Organon Teknika, Durham, N.C.), diluted 1/500 or 1/1000 in PBS/Tween 20, 100 μl of *o*-phenylenediamine dihydrochloride (OPD, Sigma) solution is added to each well and incubated for 5-15 minutes. The OPD solution is prepared by dissolving a 5 mg OPD tablet in 50 ml 1% methanol in H_2O and adding 50 μl 30% H_2O_2 immediately before use. The reaction is stopped by adding 25 l of 4M H_2SO_4 Absorbance are read at 490 nm in a microplate reader (Bio-Rad).

Example 6: Preparation of vaccines.

[0139] The present invention also relates to a method of stimulating an immune response against cells that express GPCR polypeptides in a patient using GPCR polypeptides of the invention that acts as an antigen produced by or associated with a malignant cell. This aspect of the invention provides a method of stimulating an immune response in a human against cells that express a GPCR polynucleotide and/or polypeptide of the present invention. The method comprises the step of administering to a human an immunogenic amount of a polypeptide comprising: (a) the amino acid sequence of a human GPCR protein or (b) a mutein or variant of a polypeptide comprising the amino acid sequence of a human GPCR protein.

Example 7: Generation of transgenic animals expressing polypeptides as a means for testing therapeutics.

[0140] GPCR nucleic acids are used to generate genetically modified non-human animals, or site specific gene modifications thereof, in cell lines, for the study of function or regulation of disease-related genes, or to create animal models of diseases, including Parkinson's disease, Alzheimer's disease and leukemia. The term "transgenic" is intended to encompass genetically modified animals having an exogenous GPCR gene(s) that is stably transmitted in the host cells where the gene(s) may be altered in sequence to produce a modified protein, or having an exogenous GPCR LTR promoter operably linked to a reporter gene. Transgenic animals may be made through a nucleic acid construct randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. Of interest are transgenic mammals, e.g. cows, pigs, goats, horses, etc., and particularly rodents, e.g. rats, mice, etc.

[0141] The modified cells or animals are useful in the study of GPCR gene function and regulation. For example, a series of small deletions and/or substitutions may be made in the GPCR genes to determine the role of different domains in Parkinson's disease, Alzheimer's disease and leukemia. Specific constructs of interest include, but are not limited to, anti-sense constructs to block GPCR gene expression, expression of dominant negative GPCR gene mutations, and differential expression of a GPCR gene. Expression of a GPCR gene or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development is provided. In addition, by providing expression of proteins derived from GPCR in cells in which it is otherwise not normally produced or produced at inadequate levels, changes in cellular behavior can be induced.

[0142] DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. For various techniques for transfecting mammalian cells, see Keown et al., Methods in Enzymology 185:527-537 (1990).

[0143] For embryonic stem (ES) cells, an ES cell line is employed, or embryonic cells is obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of appropriate growth factors, such as leukemia inhibiting factor (LIF). When ES cells are transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting chimeric animals screened for cells bearing the construct. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected.

[0144] The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs are maintained as allogeneic or congenic grafts or transplants, or in *in* vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals are used in functional studies, drug screening, etc., e.g. to determine the effect of a candidate drug on Parkinson's disease, Alzheimer's disease and leukemia, to test potential therapeutics or treatment regimens, etc.

[0145] All publications and patent applications mentioned in this specification are incorporated herein by reference to the same extent as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[0146] The foregoing description of preferred embodiments of the invention has been presented by way of illustration and example for purposes of clarity and understanding. It

is not intended to be exhaustive or to limit the invention to the precise forms disclosed. It will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that many changes and modifications may be made thereto without departing from the spirit of the invention. It is intended that the scope of the invention be defined by the appended claims and their equivalents.

CLAIMS

We claim:

1. A method for identifying a gene whose expression level is associated with a disease state, the method comprising:

identifying at least one gene having a nucleic acid sequence encoding a protein comprising a physical characteristic;

selecting a polynucleotide sequence from the nucleic acid sequence, wherein the polynucleotide sequence is specific for a protein comprising the physical characteristic;

detecting a level of expression of the polynucleotide sequence or a complement thereof in a diseased tissue sample;

detecting a level of expression of the polynucleotide sequence or a complement thereof in a normal tissue sample; and

comparing the level of expression of the polynucleotide sequence or a complement in the diseased tissue sample to a level of expression of the gene in the control tissue sample, wherein an altered level of expression of the polynucleotide sequence or a complement in the diseased tissue sample correlates with the disease state.

2. The method of claim 1, wherein the physical characteristic comprises seven transmembrane (7-TM) domains.

3. The method of claim 2, wherein the gene is a novel G-protein linked cell receptor type gene.

4. The method of claim 1, wherein the physical characteristic comprises an amino acid sequence comprising an Asp-Arg-Tyr (DRY) motif.

5. The method of claim 4, wherein the gene is a novel G-protein linked cell receptor type gene.

6. The method of claim 1, wherein the physical characteristic comprises a signal peptide sequence characteristic of a secreted protein.

7. The method of claim 1, wherein the physical characteristic comprises a signal peptide sequence characteristic of a mitochondrial protein.

8. The method of claim 1, wherein the physical characteristic comprises an amino acid sequence characteristic of a structural feature of the protein.

9. The method of claim 1, wherein the physical characteristic comprises an amino acid sequence characteristic of a function of the protein.

10. The method of claim 1, wherein the identification of the gene comprises searching a nucleic acid sequence database for nucleic acid sequences which encode a protein comprising the physical characteristic.

11. The method of claim 10, wherein the nucleic acid sequence database is an electronic library.

12. The method of claim 10, wherein the gene is identified using a search algorithm.

13. The method of claim 1, wherein the identification of the gene comprises selecting at least one gene whose expression is known to correlate with a disease state.

14. The method of claim 1, wherein the detection of the level of expression of the polynucleotide sequence comprises:

selecting at least one isolated oligonucleotide comprising the polynucleotide sequence or a fragment thereof;

contacting the oligonucleotide with a nucleic acid preparation from the tissue sample; and

detecting a level of expression of the polynucleotide sequence by detecting an amount of hybridization of the nucleic acid preparation to the oligonucleotide under stringent conditions.

15. The method of claim 14, wherein the oligonucleotide is attached to a solid support.

16. The method of claim 15, wherein the solid support is a microarray.

17. The method of claim 14, wherein the selection of the polynucleotide sequence comprises determining at least one of a set of factors comprising (i) a redundancy of the sequence, (ii) an efficiency of hybridization to a complementary sequence, and (iii) a likelihood of the polynucleotide sequence comprising an intron.

18. The method of claim 14, wherein the nucleic acid preparation from the tissue sample comprises a detectable label.

19. The method of claim 18, wherein the detectable label is selected from the group consisting of a fluorescent label, an enzymatic label, a chemiluminescent label, a colorimetric label, and a radioactive label.

20. The method of claim 14, wherein the nucleic acid preparation from the tissue sample is amplified before detection.

21. The method of claim 20, wherein the amplification is a conducted by a polymerase chain reaction (PCR).

22. The method of claim 20, wherein the amplification is a conducted by a quantitative polymerase chain reaction (QPCR).

23. The method of claim 1, wherein comparing the level of expression of the gene comprises:

providing at least one isolated oligonucleotide comprising the polynucleotide sequence or a fragment thereof;

contacting the oligonucleotide with an amount of nucleic acid preparation from a disease tissue sample;

contacting the oligonucleotide with an equal amount of nucleic acid preparation from a normal tissue sample; and

comparing the level of expression of the polynucleotide sequence in the tissue samples by detecting an amount of hybridization of each nucleic acid preparation to the oligonucleotide under stringent conditions.

24. The method of claim 23, wherein the polynucleotide is attached to a solid support.

25. The method of claim 24, wherein the solid support is a microarray.

26. The method of claim 23, wherein the nucleic acid preparation is an RNA preparation.

27. The method of claim 26, wherein the RNA preparation is further processed to generate a labeled nucleic acid probe.

28. The method of claim 27, wherein the labeled nucleic acid probe comprises a label coupled to the probe, wherein the label is selected from the group consisting of a biotin, an avidin, a streptavidin, an antibody, an antigen, a peptide, a fluorescent label, an enzymatic label, a chemiluminescent label, a colorimetric label, and a radioactive label.

29. A method for detecting an expression of a gene identified by the method of claim 1 in a cell, the method comprising:

cloning a polynucleotide fragment comprising a sequence of the cloned gene in an expression vector; and

detecting a corresponding protein in a cell transformed with the vector comprising the cloned fragment.

30. The method of claim 29, wherein the protein is detected by an antibody.

31. The method of claim 29, wherein the protein is detected by a monoclonal antibody.

32. A method for preparing an antibody specific for a polypeptide product of a gene identified by the method of claim 1, the method comprising:

cloning a polynucleotide fragment comprising a sequence of the cloned gene in an expression vector;

isolating a polypeptide expressed by the vector, wherein the polypeptide comprises an amino acid sequence corresponding to the cloned polynucleotide;

immunizing an animal with the isolated polypeptide; and

isolating anti-peptide antibodies specific for the isolated polypeptide from the immunized animal.

33. An isolated novel polynucleotide comprising a gene whose expression level is associated with a disease state, the polynucleotide comprising a nucleic acid sequence encoding a protein comprising a physical characteristic, wherein the polynucleotide or a fragment thereof is differentially expressed in a diseased tissue sample as compared to a normal tissue sample.

34. The isolated polynucleotide of claim 33, comprising a nucleic acid sequence which encodes a protein comprising at least one of the characteristics of: (a) seven transmembrane (7-TM) domains, (b) an amino acid sequence comprising an Asp-Arg-Tyr (DRY) motif and (c) a signal peptide.

35. The isolated polynucleotide of claim 33, wherein the polynucleotide comprises a detectable label.

36. The isolated polynucleotide of claim 33, wherein the polynucleotide, or fragment thereof, is attached to a solid support.

37. The isolated polynucleotide of claim 33, wherein the polynucleotide is single stranded.

38. The isolated polynucleotide of claim 33, wherein the polynucleotide is double stranded.

39. A host cell, comprising the isolated polynucleotide of claim 33.

40. An array comprising at least two polynucleotides according to claim 33.

41. A composition, comprising a test cell sample and an isolated polynucleotide according to claim 33.

42. An electronic library comprising at least one isolated polynucleotide according to claim 33.

43. An isolated polynucleotide comprising:

(a) a polynucleotide having sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement;

(b) a fragment of the polynucleotide having the sequence shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5), or its complement, wherein the fragment is at least 10 nucleotides in length; or

(c) a polynucleotide that selectively hybridizes to the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5) or the fragment in (b),

wherein expression of the isolated polynucleotide correlates to a state of disease.

44. The isolated polynucleotide of claim 43, comprising a segment of up to at least 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000 or 1500 nucleotides in length which corresponds identically to a portion of the sequences shown in Figures 1B, 2B, 3B, 4B and 5B (SEQ ID NOS 1-5).

45. The isolated polynucleotide of claim 43, wherein the polynucleotide comprises a detectable label.

46. The isolated polynucleotide of claim 43, wherein the polynucleotide is attached to a solid support.

47. The isolated polynucleotide of claim 43, wherein the polynucleotide is single stranded.

48. The isolated polynucleotide of claim 43, wherein the polynucleotide is double stranded.

49. The isolated polynucleotide of claim 43, wherein expression levels of the polypeptide correlate to a state of Parkinson's disease, Alzheimer's disease, or leukemia.

50. A host cell, comprising the isolated polynucleotide of claim 43.

51. An array comprising at least two polynucleotides according to claim 43.

52. A composition, comprising a test cell sample and an isolated polynucleotide according to claim 43.

53. A kit for diagnosing a disease in a test sample, the kit comprising at least one isolated polynucleotide according to claim 43.

54. An electronic library comprising at least one isolated polynucleotide according to claim 43.

55. An isolated polypeptide, or a fragment thereof, whose expression levels in a tissue correlates to a disease state of the tissue, wherein the polypeptide comprises the amino acid sequence encoded by the polypeptides shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10), or a fragment thereof.

56. The isolated polypeptide of claim 55, wherein the disease state is selected from the group consisting of central nervous system (CNS) disease, Parkinson's disease, Alzheimer's disease and leukemia.

57. The isolated polypeptide of claim 55, wherein expression levels of the polypeptide correlate to a state of CNS disease, Parkinson's disease, Alzheimer's disease, or leukemia.

58. The isolated polypeptide of claim 55, wherein the polypeptide comprises a fragment that includes an antigenic epitope comprising the amino acid sequence shown in Figures 1C, 2C, 3C, 4C and 5C (SEQ ID NOS 6-10).

59. The isolated polypeptide of claim 55, wherein the polypeptide or fragment thereof, is attached to a solid support.

60. An isolated antibody, or antigen binding fragments thereof, that bind to the polypeptide according to any one of claims 55.

61. The isolated antibody of claim 60, wherein the antibody is a polyclonal antibody.

62. The isolated antibody of claim 60, wherein the antibody is a monoclonal antibody.

63. The isolated antibody of claim 60, wherein the antibody is attached to a solid surface.

64. The isolated antibody of claim 60, wherein the antibody comprises a detectable label.

Category: **GPCR-novel**

ID: gi6863021_GS_nt7

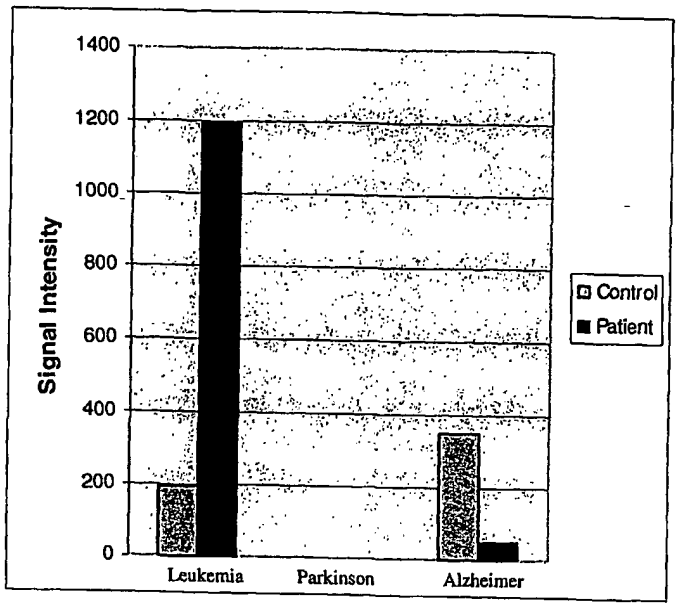

Figure 1A

Sequence:

>gi6863021_GS_nt7 AP001112

atgctcaatttcaccgatgtgacagagttcattcttttggggctaacgagccgtcgagaatggcaagtt
ctcttcttcatcatctttcttgtggtctacatcatcaccatggtgggcaatatcggcatgatggtgtta
atcaaggtcagtcctcagcttaacaaccccatgtacttttttcctcagtcacttgtcatttgttgatgtg
tggttttcttccaatgtcacccctaaaatgttggaaaacctgttttcagataaaaaaacaattacttat
gctggttgtttagtacagtgtttcttcttcattgctcttgtccatgtggaaatttttattcttgctgcg
atggcctttgatagatacatggcaattgggaatcctctgctttatggcagtaaaatgtcaagggttgtc
tgtattcgactgattactttcccttacatttatggttttctgacgagtctggcagcaacattatggact
tacggcttgtacttctgtggaaaaattgagatcaaccatttctactgtgcagatccacctctcatcaaa
atggcctgtgccgggacctttgtaaaagaatatacaatgatcatacttgccggcattaacttcacatat
tccctgactgtaattatcatctcttacttattcatcctcattgccattctgcgaatgcgctcagcagaa
ggaaggcagaaggccttttccacatgtgggtcccatctgacagctgtcattatattctatggtactctg
atcttcatgtatctcagacgtcccacagaggagtctgtggagcaggggaagatggtggctgtgttctat
accacagtgatccccatgttgaatcccatgatctacagtctgaggaacaaggatgtgaaaaaggccatg
atgaaagtgatcagcagatcatgttaa

Figure 1B

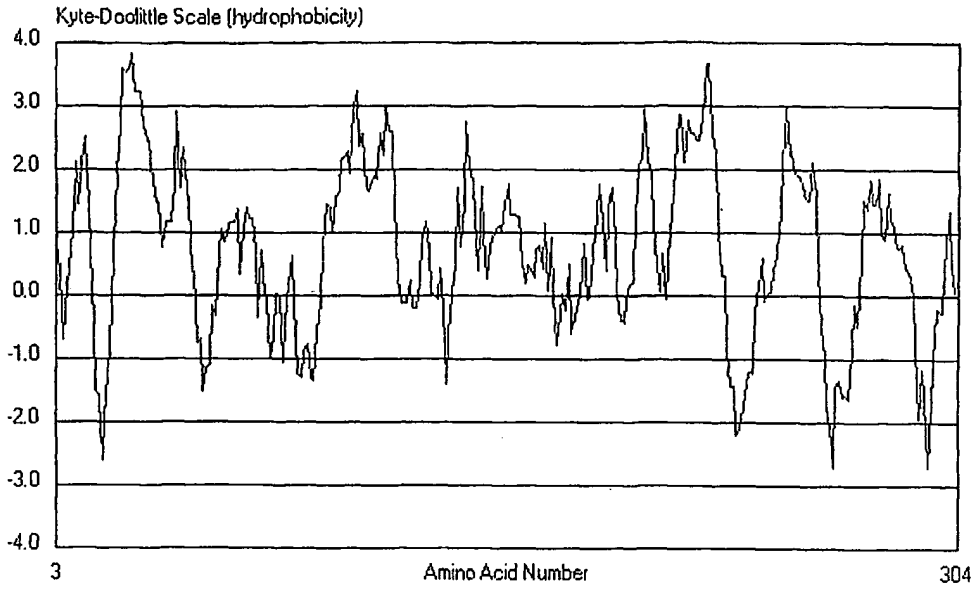

Figure 1D

>gi6863021_GS_nt7 AP001112
MLNFTDVTEFILLGLTSRREWQVLFFIIFLVVYIITMVGNIGMMVLIKVS
PQLNNPMYFFLSHLSFVDVWFSSNVTPKMLENLFSDKKTITYAGCLVQCF
FFIALVHVEIFILAAMAFDRYMAIGNPLLYGSKMSRVVCIRLITFPYIYG
FLTSLAATLWTYGLYFCGKIEINHFYCADPPLIKMACAGTFVKEYTMIIL
AGINFTYSLTVIIISYLFILIAILRMRSAEGRQKAFSTCGSHLTAVIIFY
GTLIFMYLRRPTEESVEQGKMVAVFYTTVIPMLNPMIYSLRNKDVKKAMM
KVISRSC*

Figure 1C

Category: **GPCR-N vel**

ID: gi6453999_GS_nt6

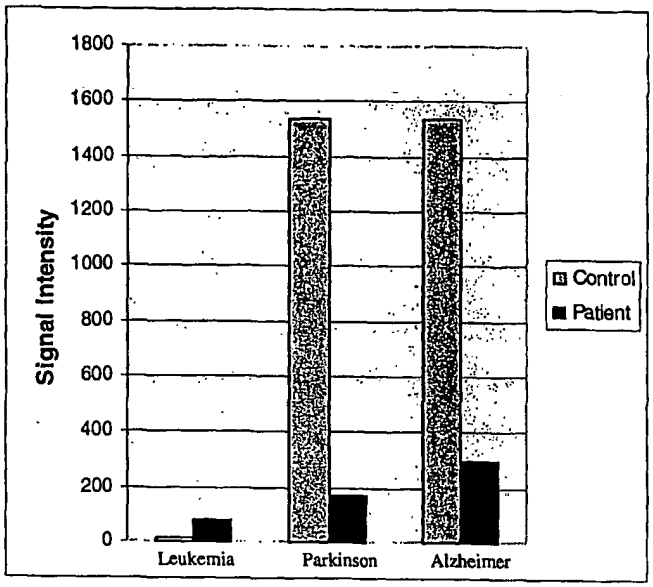

Figure 2A

Sequence:

>gi6453999_GS_nt6 AC011571

atggatcagaaaaatggaagttctttcactggatttatcctactgggtttctctg
acaggcctcagctggagctagtcctctttgtggttcttttgatcttctatatctt
cactttgctggggaacaaaaccatcattgtattatctcacttggacccacatctt
cacactcctatgtatttttcttctccaacctaagctttttggatctgtgttaca
caaccggcattgttccacagctcctggttaatctcaggggagcagacaaatcaat
ctcctatggtggttgtgtagttcagctgtacatctctctaggcttgggatctaca
gaatgcgttctcttaggagtgatggtatttgaccgctatgcagctgtttgcaggc
ccctccactacacagtagtcatgcacccttgtctgtatgtgctgatggcttctac
ttcatgggtcattggttttgccaactccctattgcagacggtgctcatcttgctt
ttaacactttgtggaagaaataaattagaacactttctttgtgaggttcctccat
tgctcaagcttgcctgtgttgacactactatgaatgaatctgaactcttctttgt
cagtgtcattattcttcttgtacctgttgcattaatcatattctcctatagtcag
attgtcagggcagtcatgaggataaagttagcaacagggcagagaaaagtgtttg
ggacatgtggctcccacctcacagtggtttccctgttctacggcacagctatcta
tgcttacctccagcccggcaacaactactctcaggatcagggcaagttcatctct
ctcttctacaccatcattacacccatgatcaacccccctcatatatacactgagga
acaaggatgtgaaaggagcacttaagaaggtgctctggaagaactacgactccag
atga

Figure 2B

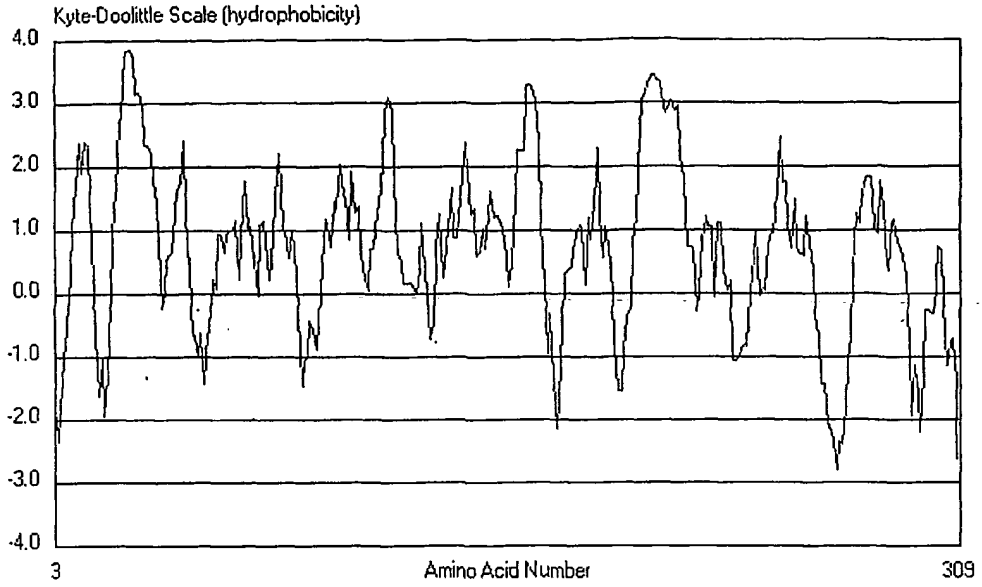

Figure 2D

>gi6453999_GS_nt6 AC011571
MDQKNGSSFTGFILLGFSDRPQLELVLFVVLLIFYIFTLLGNKTIIVLSH
LDPHLHTPMYFFFSNLSFLDLCYTTGIVPQLLVNLRGADKSISYGGCVVQ
LYISLGLGSTECVLLGVMVFDRYAAVCRPLHYTVVMHPCLYVLMASTSWV
IGFANSLLQTVLILLLTLCGRNKLEHFLCEVPPLLKLACVDTTMNESELF
FVSVIILLVPVALIIFSYSQIVRAVMRIKLATGQRKVFGTCGSHLTVVSL
FYGTAIYAYLQPGNNYSQDQGKFISLFYTIITPMINPLIYTLRNKDVKGA
LKKVLWKNYDSR*

Figure 2C

Category: **GPCR-novel**

ID: gi6671985_GS_nt9.2

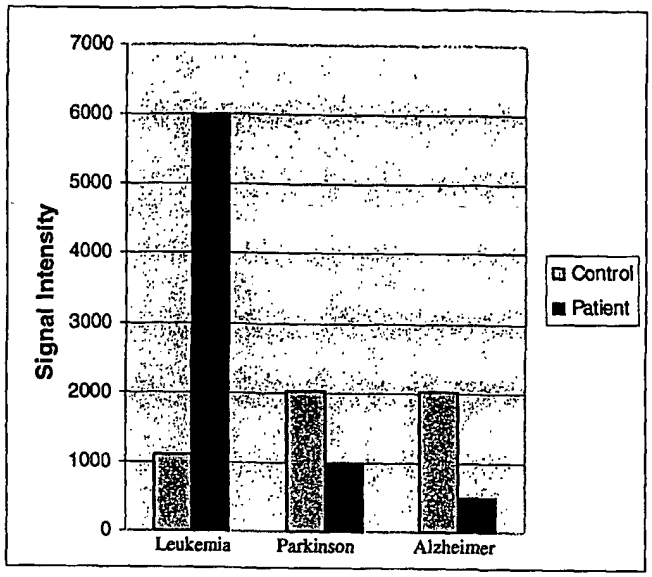

Figure 3A

Sequence:

>gi6671985_GS_nt9.2

atggctgcaggaaatcactctacagtgacagagttcattctcaagggtttaacgaagagagcag
acctccagctccccctctttctcctcttcctcgggatctacttggtcaccatcgtggggaacct
gggcatgatcactctaatttgtctgaactctcagctgcacacccccatgtactactttctcagc
aatctgtcactcatggatctctgctactcctccgtcattacccctaagatgctggtgaactttg
tgtcagagaaaaacatcatctcctacgcagggtgcatgtcacagctctacttcttccttgtttt
tgtcattgctgagtgttacatgctgacagtgatggcctacgaccgctatgttgncntctgccac
cctttgctttacaacatcattatgtctcatcacacctgcctgctgctggtggctgtggtctacg
ccatcggactcattggctccacaatagaaactggcctcatgttaaaactgccctattgtgagca
cctcatcagtcactacttctgtgacatcctccctctcatgaagctgtcctgctctagcacctat
gatgttgagatgacagtcttcttttcggctggattcaacatcatagtcacgagcttaacagttc
ttgtttcttacaccttcattctctccagcatcctcggcatcagcaccacagaggggagatccaa
agccttcagcacctgcagctcccaccttgcagccgtgggaatgttctatggatcaactgcattc
atgtacttaaaaccctccacaatcagttccttgacccaggagaatgtggcctctgtgttctaca
ccacggtaatccccatgttgaatcccctaatctacagcctgaggaacaaggaagtaaaggctgc
cgtgcagaaaacgctgaggggtaaactgttttga

Figure 3B

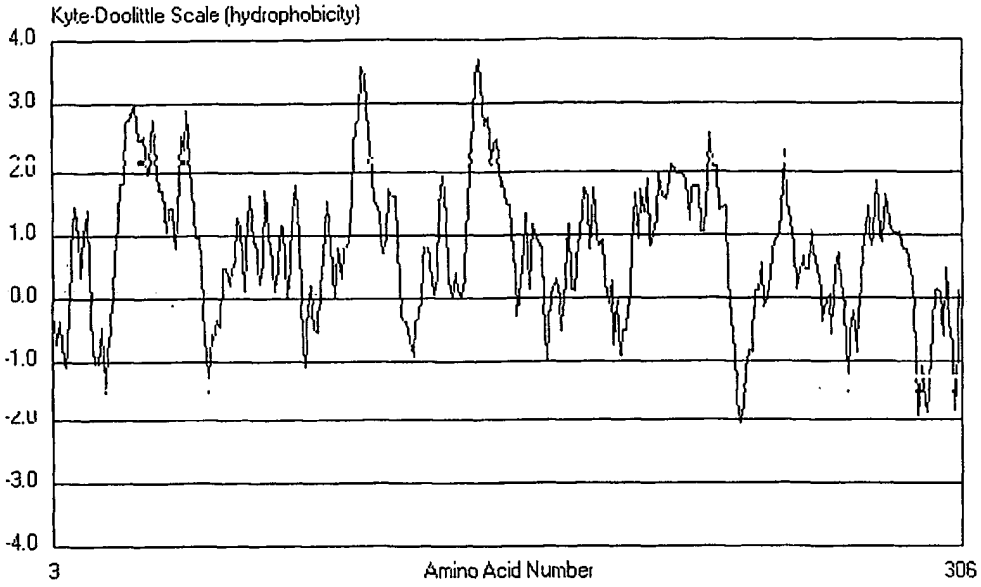

Figure 3D

>gi6671985_GS_nt9.2

MAAGNHSTVTEFILKGLTKRADLQLPLFLLFLGIYLVTIVGNLGMITLIC
LNSQLHTPMYYFLSNLSLMDLCYSSVITPKMLVNFVSEKNIISYAGCMSQ
LYFFLVFVIAECYMLTVMAYDRYVXXCHPLLYNIIMSHHTCLLLVAVVYA
IGLIGSTIETGLMLKLPYCEHLISHYFCDILPLMKLSCSSTYDVEMTVFF
SAGFNIIVTSLTVLVSYTFILSSILGISTTEGRSKAFSTCSSHLAAVGMF
YGSTAFMYLKPSTISSLTQENVASVFYTTVIPMLNPLIYSLRNKEVKAAV
QKTLRGKLF*

Figure 3C

Category: **GPCR-Novel**

ID: gi5791525_GS_nt10

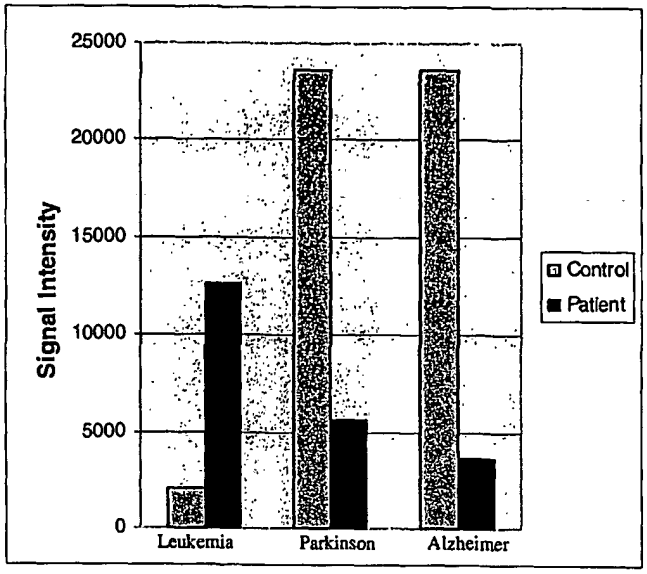

Figure 4A

Sequence:

>gi5791525_GS_nt10 HSJ154J13

atggccatggacaatgtcacagcagtgtttcagtttctccttattggcatttctaactatcct
caatggagagacacgtttttcacattagtgctgataatttacctcagcacattgttggggaat
ggatttatgatctttcttattcactttgaccccaacctccacactccaatctacttcttcctt
agtaacctgtctttcttagacctttgttatggaacagcttccatgccccaggctttggtgcat
tgtttctctacccatccctacctctcttatccccgatgtttggctcaaacgagtgtctccttg
gctttggccacagcagagtgcctcctactggctgccatggcctatgaccgtgtggttgctatc
agcaatcccctgcgttattcagtggttatgaatggcccagtatgtgtctgcttggttgctacc
tcatgggggacatcacttgtgctcactgccatgctcatcctatccctgaggcttcacttctgt
ggggctaatgtcatcaaccattttgcctgtgagattctctccctcattaagctgacctgttct
gataccagcctcaatgaatttatgatcctcatcaccagtatcttcaccctgctgctaccattt
gggtttgttctcctctcctacatacgaattgctatggctatcataaggattcgctcactccag
ggcaggctcaaggcctttaccacatgtggctctcacctgaccgtggtgacaatcttctatggg
tcagccatctccatgtatatgaaaactcagtccaagtcctaccctgaccaggacaagtttatc
tcagtgttttatggagctttgacacccatgttgaaccccctgatatatagcctgagaaaaaaa
gatgttaaacgggcaataaggaaagttatgttgaaaaggacatga

Figure 4B

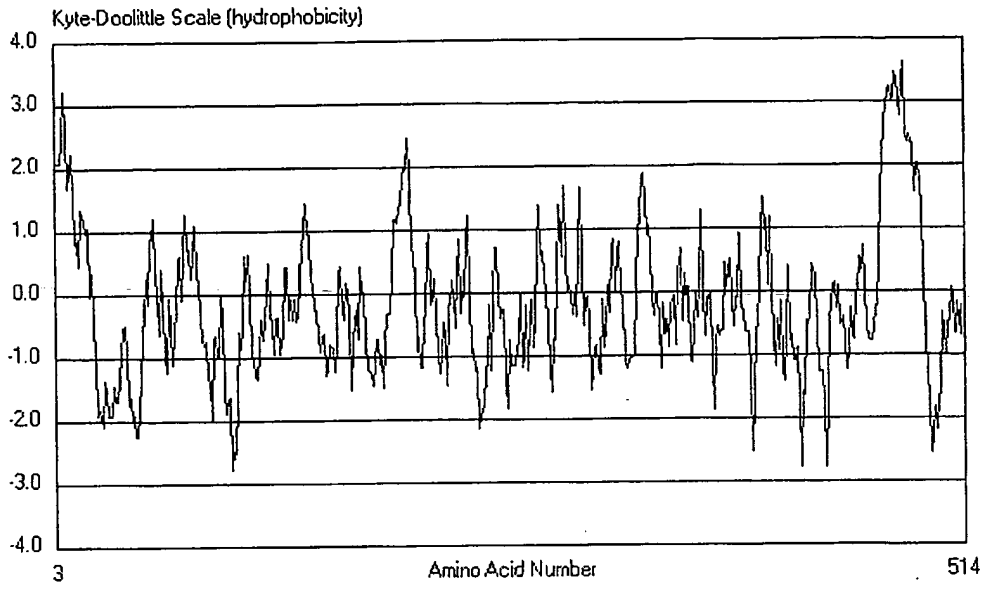

Figure 4D

>gi5791525_GS_nt10 HSJ154J13
MAMDNVTAVFQFLLIGISNYPQWRDTFFTLVLIIYLSTLLGNGFMIFLIH
FDPNLHTPIYFFLSNLSFLDLCYGTASMPQALVHCFSTHPYLSYPRCLAQ
TSVSLALATAECLLLAAMAYDRVVAISNPLRYSVVMNGPVCVCLVATSWG
TSLVLTAMLILSLRLHFCGANVINHFACEILSLIKLTCSDTSLNEFMILI
TSIFTLLLPFGFVLLSYIRIAMAIIRIRSLQGRLKAFTTCGSHLTVVTIF
YGSAISMYMKTQSKSYPDQDKFISVFYGALTPMLNPLIYSLRKKDVKRAI
RKVMLKRT*

Figure 4C

Category: **GPCR-novel**

ID: gi6454063_GS_nt8

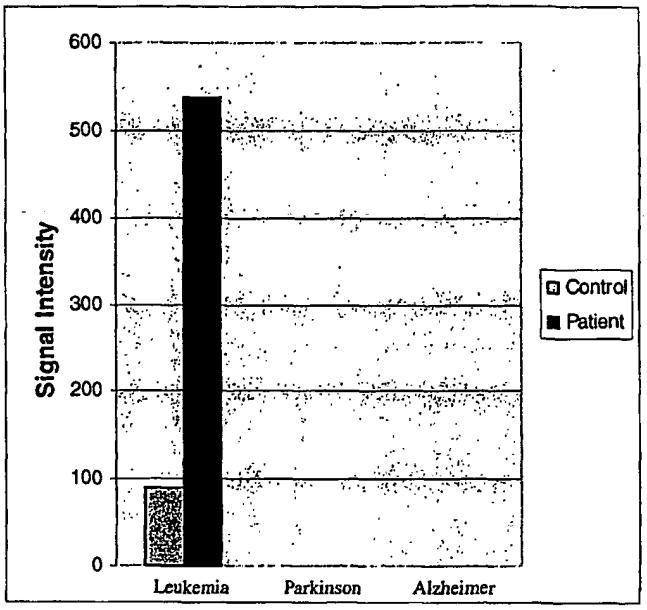

Figure 5A

Sequence:

>gi6454063_GS_nt8 AC010930

atgtcagcctccaatatcaccttaacacatccaactgccttcttgttggtggggattccaggcct
ggaacacctgcacatctggatctccatccctttctgcttagcatatacactggccctgcttggaa
actgcactctccttctcatcatccaggctgatgcagccctccatgaacccatgtacctctttctg
gccatgttggcagccatcgacctggtcctttcctcctcagcactgcccaaaatgcttgccatatt
ctggttcagggatcgggagataaacttctttgcctgtctggcccagatgttcttccttcactcct
tctccatcatggagtcagcagtgctgctggccatggcctttgaccgctatgtggctatctgcaag
ccactgcactacaccaaggtcctgactgggtccctcatcaccaagattggcatggctgctgtggc
ccgggctgtgacactaatgactccactccccttcctgctgagatgtttccactactgccgaggcc
cagtgatcgctcactgctactgtgaacacatggctgtggtgaggctggcgtgtggggacactagc
ttcaacaatatctatggcatcgctgtggccatgtttattgtggtgttggacctgctccttgttat
cctgtcttatatctttattcttcaggcagttctactgcttgcctctcaggaggcccgctacaagg
catttgggacatgtgtctctcatataggtgccatcttagccttctacacaactgtggtcatctct
tcagtcatgcaccgtgtagcccgccatgctgcccctcatgtccacatcctccttgccaatttcta
tctgctcttcccacccatggtcaatcccataatctatggtgtcaagaccaagcaaatccgtgaga
gcatcttgggagtattcccaagaaaggatatgtag

Figure 5B

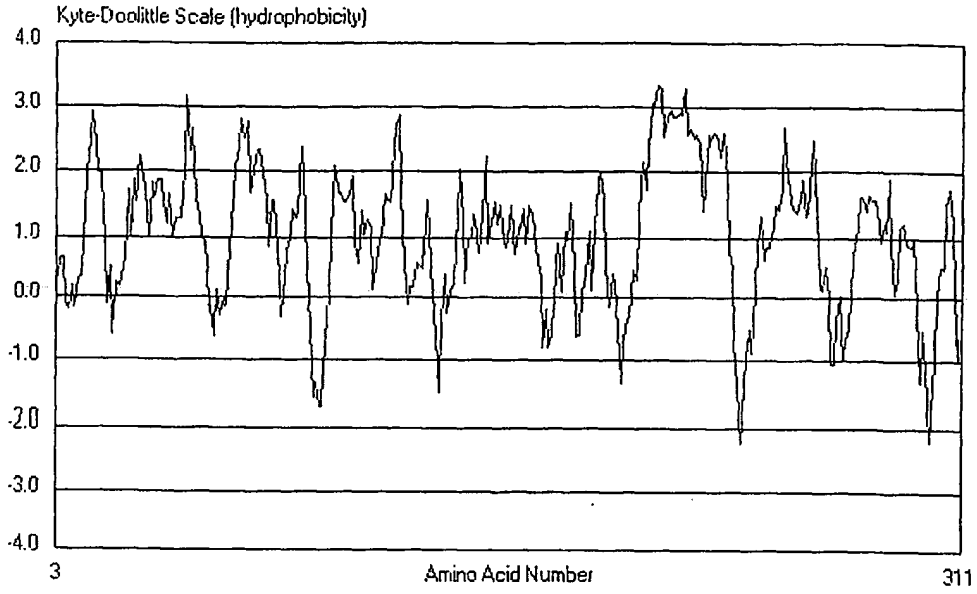

Figure 5D

>gi6454063_GS_nt8 AC010930
MSASNITLTHPTAFLLVGIPGLEHLHIWISIPFCLAYTLALLGNCTLLLI
IQADAALHEPMYLFLAMLAAIDLVLSSSALPKMLAIFWFRDREINFFACL
AQMFFLHSFSIMESAVLLAMAFDRYVAICKPLHYTKVLTGSLITKIGMAA
VARAVTLMTPLPFLLRCFHYCRGPVIAHCYCEHMAVVRLACGDTSFNNIY
GIAVAMFIVVLDLLLVILSYIFILQAVLLLASQEARYKAFGTCVSHIGAI
LAFYTTVVISSVMHRVARHAAPHVHILLANFYLLFPPMVNPIIYGVKTKQ
IRESILGVFPRKDM*

Figure 5C

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2002 (28.11.2002)

PCT

(10) International Publication Number
WO 02/095065 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/DK02/00337

(22) International Filing Date: 21 May 2002 (21.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2001 00802 18 May 2001 (18.05.2001) DK

(71) Applicant *(for all designated States except US)*: **AZIGN BIOSCIENCE A/S** [DK/DK]; Vestre Teglgade 10, DK-2450 København SV (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **THIRSTRUP, Kenneth** [DK/DK]; Frederiksborgvej 141 st.tv., DK-2400 København NV (DK). **MADSEN, Lars, Siim** [DK/DK]; Kingosgade 5, 3.th, DK-1623 København V (DK). **JENSEN, Jens, Bitsch** [DK/DK]; Trøjborggade 7, 5 tv., DK-1757 København V (DK). **HUMMEL, Rene** [DK/DK]; Ellemosevej 73, DK-2900 Hellerup (DK). **JENSEN, Bo, Skaaning** [DK/DK]; Genuavej 22 C, DK-2300 København S (DK).

(74) Agents: **HALBERG, Kristian** et al.; NeuroSearch A/S, Patent Department, 93 Pederstrupvej, DK-2750 Ballerup (DK).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/095065 A2

(54) Title: G-PROTEIN COUPLED RECEPTOR ARRAYS

(57) Abstract: The invention relates to G-protein coupled receptor (GPCR) arrays, methods for production of GPCR arrays, primers used in the production of GPCR arrays and kits containing GPCR arrays and further to the use of such GPCR arrays in methods for the determination of expression profiles in biological materials in which there is an interest in the expression of GPCR polynucleotides.

G-PROTEIN COUPLED RECEPTOR ARRAYS

FIELD OF INVENTION

The invention relates to G-protein coupled receptor (GPCR) arrays, methods for production of GPCR arrays, primers used in the production of GPCR arrays and kits containing GPCR arrays and further to the use of such GPCR arrays in methods for the determination of expression profiles in biological materials in which there is an interest in the expression of GPCR polynucleotides.

BACKGROUND OF INVENTION

Different kinds of arrays have become increasingly important tools in the biotechnology industry and related fields. These arrays have a plurality of polynucleotide spots deposited on a solid surface in form of an array. Arrays of both polypeptides and polynucleotides have been developed and find use in a variety of applications. One of the applications is differential gene expression, where expression of genes in different cells or tissues (normally a control sample and a sample of the cell or tissue of interest) is compared and any difference in the mRNA expression profile is determined.

In gene expression analysis using arrays, an array of "probe" nucleotides is contacted with a nucleic acid sample of interest such as mRNA concerted into cDNA from a particular tissue or cell. Contact is carried out under hybridisation conditions favourable for hybridisation of nucleic acids complementary to the "probe" nucleotides on the array. Unbound nucleic acid is then removed by washing. The resulting pattern of hybridised nucleic acid provides information regarding the gene expression profile of the sample tested on the array. Gene expression analysis is used in a variety of applications including identification of novel expression of genes, correlation of gene expression to a particular tissue or a particular disease, identifying effects of agents on the cellular expression such as in toxicity testing and in identifying drugs.

A variety of different array techniques have been developed during the years in order to meet the growing demands from the biotechnology industry, see e.g. *Lockhart et al.*, Nature Biotechnology 1996 **14** 1675-1680; *Shena et al.*, Science 1995 **270** 467-470 and WO 98/51789. However, there is still a need for new improved arrays having specific applications.

We hereby provide a novel array capable of determining the polynucleotide expression profile of GPCR derived polynucleotides in a biological material.

BRIEF DISCLOSURE OF THE INVENTION

The object of the invention is to provide GPCR arrays, kits comprising GPCR arrays and methods to produce such GPCR arrays. The GPCR arrays may be used for the determination of GPCR expression profiles in biological materials and also for the identification of therapeutic, prophylactic and/or toxic agents, where the therapeutic, prophylactic and/or toxic agents directly or indirectly influence the GPCR expression profiles in biological materials.

Accordingly, in a first aspect the invention relates to a GPCR array comprising a multiplicity of individual GPCR polynucleotide spots stably associated with a surface of a solid support, wherein an individual GPCR polynucleotide spot comprises a GPCR polynucleotide composition comprising a non-conserved region of a GPCR polynucleotide family member, the spots representing at least two different regions of a GPCR polynucleotide member of a family.

In another aspect, the invention relates to a method of preparing an array according to the invention, said method comprising generating said non-conserved regions of GPCR polynucleotide family members, preparing a multiplicity of compositions each comprising at least a non-conserved region, and stably associating said compositions in individual spots on a surface of a solid support.

In a further aspect, the invention relates to a set of primers specific for non-conserved regions of GPCR polynucleotide family members, wherein the set of primers are used in the method for the production of an array according to the invention.

In still a further aspect, the invention relates to a method for the determination of a GPCR polynucleotide expression profile in a biological material, said method comprising, obtaining a polynucleotide from the biological material, labelling said polynucleotide to obtain a labelled target polynucleotide sample, contacting at least one labelled target polynucleotide sample with an array according to the invention under conditions which are sufficient to produce a hybridisation pattern, and detecting said hybridisation pattern to obtain the GPCR polynucleotide expression profile of the biological material.

In still a further aspect, the invention relates to a method for the determination of a difference in GPCR polynucleotide expression profiles from at least a first and a second different biological materials, said method comprising obtaining a first GPCR expression profile of the first biological material according to the method of the present invention obtaining a second GPCR expression profile of the second biological material according to the method of the present invention, comparing the

first and the second GPCR expression profile to identify any differences in the GPCR expression profiles between the first and the second GPCR expression profile.

In still a further aspect, the invention relates to a method for identifying a therapeutic, prophylactic and/or toxic agent involved in the response of GPCR polypeptides in a biological material, said method comprises obtaining a first GPCR expression profile of a first biological material according to the method of the present invention, obtaining a second GPCR expression profile of a second biological material according to the method of the present invention, treating the second biological material with a test compound; obtaining a third GPCR expression profile of the treated second biological material according to the method of the present invention, comparing the first, second and third GPCR expression profiles, and identifying any difference in the GPCR expression profile so as to identify any therapeutic, prophylactic or toxic response of the test compound on the GPCR polynucleotide in the second biological material.

In still a further aspect, the invention relates to a diagnostic method to determine the differences of GPCR expression profiles between two different biological materials; said method comprises obtaining a first GPCR expression profile of a first biological material according to the method of the present invention, obtaining a second GPCR expression profile of a second biological material according to the method of the present invention, comparing the first and second GPCR expression profiles, and identifying any difference in the GPCR expression profile.

In a final aspect, the invention relates to a GPCR kit for use in a hybridisation assay, said kit comprising a GPCR array according to the present invention.

The invention provides novel and improved GPCR arrays, kits comprising GPCR arrays and methods to produce such GPCR arrays. GPCR arrays are useful in the determination of GPCR expression profiles in biological materials and also in the identification of therapeutic, prophylactic and/or toxic agents; the therapeutic, prophylactic and/or toxic agent may directly or indirectly influence the GPCR expression profiles in biological materials.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a novel GPCR array. The GPCR array comprises e.g. a slide or a membrane or other kind of solid support onto which polynucleotide spots are applied and the polynucleotide spots represent one or more GPCR families or family members. Preferably, the polynucleotides (or fragments of polynucleotides) spotted on the slides have been chosen in such a way that the

polynucleotides have specificity for more than one species such as e.g. humans, rats and mice, i.e. there is a certain interspecies identity, and at the same time, the polynucleotides chosen from a GPCR family member preferably have a certain degree of non-identity with other family members belonging to the same GPCR family, i.e. there is a relatively low intrafamily identity. By choosing the polynucleotides in such a manner, the use of a GPCR array of the invention makes it possible to obtain information relating to a GPCR family in general (and not only to a specific GPCR family member) and at the same time it is possible to compare and utilize information relevant for different species, cf. below.

Having a GPCR array containing polynucleotides spots with specificity for e.g. human, rat and mouse makes it possible to use the GPCR array on both human, rat and mouse derived biological material. It is therefore envisaged that the use of a GPCR array according to the present invention will lead to a better understanding of e.g. the pathogenesis of different ion related conditions or diseases in humans, since it is possible to compare biological material from humans with relevant biological material from e.g. well-known disease models in e.g. rats and mice.

If a GPCR is contemplated to be involved in the pathogenesis of a human related disease (e.g. as evidenced by the results of an analysis of RNA extracted from biopsies) the present invention makes it possible to confirm such a hypothesis by employing a GPCR array according to the invention. In such a case disease models in e.g. rats and mice are used and biological material obtained from e.g. diseased and healthy rats and/or mice is assayed by means of a GPCR array of the invention. In the same manner it is possible to confirm a hypothesis that a GPCR is involved in the pathogenesis in a disease model in e.g. rat or mouse (e.g. evidenced by analysis of RNA extracted from relevant tissue) also is associated with the condition in human. In such a case, RNA extracted from e.g. human biopsies is assayed by means of a GPCR array according to the invention and compared with the results from assays employing the relevant biological material from the diseased (and healthy) animals. Furthermore, it is possible to assess the expression level of the GPCR's across different tissues. In that way it is possible to address the question of the specificity the GPCR in the particular disease and gain information about the suitability of the GPCR as a target for the disease model.

As mentioned above, the spotted polynucleotides on the array of the invention have a certain cross-specificity to e.g. both human, rat and mouse. Samples from other species (e.g. pigs, dogs, chickens, cows and the like) can also be analysed on the GPCR array. This makes it possible to analyse tissue from other species with an unknown disease and compare it with tissue from well known e.g. neurological as well as other disease models in mice and rate. Such an approach leads to a better

understanding of diseases in other species and initiates therapeutic strategies for e.g. economically expensive disease present in e.g. animal households.

Definitions

In the context of the present application and invention the following definitions apply:

The term "polynucleotide" is intended to mean a single or double stranded polymer composed of nucleotides, e.g. deoxyribonucleotides and/or ribonucleotides from about 30 to about 9,000 nucleotides in length, from about 50 to about 6,000, from about 50 to about 3,000, from about 50 to about 1,500, from about 50 to about 1,000, from about 100 to about 1,000, from about 200 to about 750, from about 200 to 700, from about 200 to 500 or from about 300 to about 350. The polynucleotides may be single or double stranded polynucleotides.

The term "complementary" or "complementarity" is used in relation to the base-pairing rules of nucleotides well known for a person skilled in the art. Polynucleotides may be complete or partial complementary. Partial complementarity means that at least one nucleic acid base is not matched according to the base pairing rules. Complete complementarity means that all nucleotides in a polynucleotide match according to the base pairing rules. The degree of complementary between polynucleotides affects the strength of hybridisation between two polynucleotide strands. The inhibition by hybridisation of the complementary polynucleotide to the target polynucleotide may be analysed by techniques well known for a person skilled in the art, such as Southern blot, Northern blot, and the like under conditions of high stringency. A partially (substantially) homologous polynucleotide will compete for and inhibit the binding of a completely homologous sequence to the target sequence under low stringency.

The term "homology" is intended to mean the degree of identity of one polynucleotide to another polynucleotide. According to the invention the term homology is used in connection with complementarity between polynucleotides within a family or between species. There may be complete homology (i.e. 100% identity) between two or more polynucleotides. The degree of homology may be determined by any method well known for a person skilled in the art.

The term "polynucleotide composition" is intended to mean a composition comprising a polynucleotide together with an excipient. The polynucleotide compositions are applied as spots on the array. The GPCR polynucleotide composition comprises a non-conserved region of a GPCR polynucleotide family member. The term "polynucleotide composition" includes also control or calibrating

compositions such as, e.g. compositions comprising polynucleotides corresponding to housekeeping genes.

The term "non-conserved region" is intended to mean a segment of nucleotides in a polynucleotide, which compared to a segment of nucleotides in another polynucleotide has at the most about 90% identity. A non-conserved region of a GPCR polynucleotide family member is thus defined as a region of nucleotides corresponding to part of the polynucleotide, and the non-conserved region has less than 90% such as, e.g. less than 85% less than about 80%, less than about 75% or less than about 70% identity compared to all other polynucleotides belonging to the same GPCR polynucleotide family (intrafamily identity). In particular the non-conserved regions will be found in polynucleotides corresponding to the intra- or extra cellular loops or the N- or C-terminal part of GPCR.

Accordingly, the term "conserved region" is intended to mean a segment of nucleotides in a polynucleotide, which compared to a segment of nucleotides in another polynucleotide has more than 90%, such as at least about 92%, at least about 95% or at least about 97% identity.

The term "GPCR" is intended to mean a polypeptide which is a transmembrane protein consisting of seven transmembrane domains with same overall structure. The majority of the GPCR's have the ability to transduce a signal across the cell membrane through activation of a G-protein on the intracellular side of the cell. However, some of these receptors may also be signalling via alternative signal molecules like Jak2 kinases, phospholipase Cγ or protein kinase C.

The GPCR's are classified upon their ligand specificity, mode of activation, biological function, regulation or molecular structure. More than 2000 G protein coupled receptors have been reported since the cloning of bovine opsin receptor in 1983 and they have been classified into over 100 subfamilies according to sequence homology, ligand structure, and receptor function. Examples of GPCR's are adrenergic, adenosine, dopaminergic, histaminergic, opioid and serotonin GPCR's.

GPCR's are direct or indirect targets for the action of compounds, such as drugs.

The term "GPCR polynucleotide" is intended to mean a polynucleotide encoding a polypeptide (GPCR) involved in transducing a signal across biological membranes.

The term "GPCR family" is intended to mean a group of GPCR polypeptides, which has common characteristics such as, e.g. ligand specificity (i.e. the ability to bind a ligand with almost similar affinity and efficacy), homology, and same overall tertiary amino acid structure. Each family comprises individual members each having structural variations but they fulfil the requirements mentioned above with

respect to being classified as a family. An example of a GPCR subfamily is the opioid receptor family consisting of the μ, δ and κ-opioid receptors (see also Example 1).

The term "GPCR polynucleotide family" is intended to mean polynucleotides encoding polypeptides of a GPCR family". The polynucleotides may generally be found and downloaded from Genbank or EMBL (see e.g. http:www.ncbi.nih.org).

The term "intrafamily identity" is intended to mean identity within a group of members belonging to the same level in the taxonomic system, e.g. superfamily, family or subfamily or a number of subfamilies. The level in the taxonomic system to be used is selected depending i.a. on the number of members in the level. Preferably, the level in the taxonomic system is selected so that the number of members is from 2-20, more preferably 3-15, more preferably 4-12, and most preferably 5-10.

The term "interspecies identity" is intended to mean identity between a group of different species, such as a group comprising humans, mice and rats.

The terms "expression profile", "differential expression profile" and "gene expression profile" are intended to mean the expression of the mRNA's in a biological material. While an expression profile encompasses a representation of the expression level of at least one mRNA, in practice the typical expression profile represents the expression of several mRNA's. For example, an expression profile used according to the present invention represents the expression levels of at least from about 1 to 50,000 or more different mRNA's in a biological material. The expression level of the different mRNA's is the same or different. The expression of mRNA's may be up- or down regulated resulting in different expression profiles.

The term "biological material" includes within its meaning organisms, organs, tissues, cells or biological material produced by a cell culture. The biological material may be living or dead. The material may correspond to one or more cells from the organisms, in case the organism is a multicellular organism, the material may correspond to one or more cells from one or more tissues creating the multicellular organism. The biological material to be used according to the invention may be derived from particular organs or tissues of the multicellular organism, or from isolated cells obtained from a single or multicellular organism. In obtaining the sample of RNA's to be analysed from the biological material from which it is derived, the biological material may be subject to a number of different processing steps. Such steps might include tissue homogenisation, cell isolation and cytoplasma extraction, nucleic acid extraction and the like and such processing steps are generally well known for a person skilled in the art. Methods of isolating RNA from cells, tissues, organs or whole organisms are known to those skilled in the art and are described in e.g. *Sambrook et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbour

Press (1989). The biological material may be of the same kind i.e. the biological material is of the same kind of origin, such as coming from the same type of tissue, the same organism or the same type of organism or the same cell type etc.

The term "organism" is intended to mean any single cell organism such as yeast or multicellular organism, including plants, fungi and animals, preferably mammals, such as humans, rats, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, sheep, apes, monkeys and cats.

The term "tissue" is intended to mean a collection of differentiated cells such as adrenal gland, total brain, liver, heart, kidney, lung, pancreas, mammary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, and uterus.

The term "target polynucleotide" is intended to mean a polynucleotide present in the biological material of interest. The target polynucleotide encodes a polypeptide, which is at least a part of a GPCR. If the target polynucleotide has a complementary polynucleotide present on the GPCR array, it will hybridise thereto and thus give rise to a detectable signal.

The term "non-overlapping" is intended to mean that when the GPCR polynucleotide regions used in the GPCR polynucleotide composition spots are obtained from the same polynucleotide, the regions are obtained from different parts of the polynucleotide and the different parts are located in such a manner that the regions not even overlap each other by a single nucleotide. In a polynucleotide of e.g. 1,000 nucleotides the regions 1-500 and 501-900 are non-overlapping. The non-overlapping GPCR polynucleotide regions may be located with a distance of one or more nucleotides from each other.

The term "primer" is intended to mean a polymer of 10-50 nucleotides.

The term "set of primers" is intended to mean one or more primers having the ability to amplify a GPCR polynucleotide region under suitable conditions. The length of the primers may be the same or different and dependent on the character of the GPCR polynucleotide region to be amplified. Design of such a set of primers is well known for a person skilled in the art. The set of primers having a sufficient length to specifically hybridise to a distinct GPCR polynucleotide in the sample and the length of the primers will be from about 3 to 50 nucleotides.

The terms "stressed state" and "stressed" are intended to mean that the above described "biological material" is influenced compared to the normal condition. When an expression profile is obtained from a stressed biological material it is different compared to a non-stressed biological material. The biological material may be influenced by some kind of organic/inorganic compound, an environmental agent, a drug substance, pathogen, mutagen, mitogen, receptor mediated signal or the like.

Normally, the biological material is influenced in such a manner that the expression profile of the GPCR polynucleotides in the biological material either directly or indirectly is affected resulting in at least one difference between the expression profile of the non-stressed biological material compared to the stressed biological material.

Methods of comparing the homology between different polynucleotides and/or parts of different polynucleotides irrespective of whether the parts are conserved or non-conserved regions are well known in the art. The polynucleotides may either belong to the same family or different families and/or being polynucleotides encoding the same polypeptide from the same or different species. Optimal alignment of nucleotides of a polynucleotide for comparison of the homologies may be conducted using the homology algorithm described by e.g. *Smith and Waterman*, Adv. Appl. Math. 1981 **2** 482, by the homology alignment algorithm of *Needleman and Wunsch*, J. Mol. Biol. 1970 **48** 443, by the search for similarity method of *Pearson and Lipman*, Proc. Natl. Acad. Sci. USA 1988 **85** 2444, or by computerised implementations of these algorithms by using for example CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG) 575 Science Dr., Madison, Wisconsin, USA. The above-mentioned algorithms and computer implementations should be regarded as examples and the invention not limited thereto.

GPCRs

It has been estimated that up to 5000 distinct GPCR-encoding genes exist within the human genome. Up to now approximately 800 genes have been cloned from various species, including approximately 150 human genes. The GPCR superfamily may be classified in three major homology families for the mammalian GPCRs, viz. the family 1 or rho-family (prototype: rhodopsin), the family 2 or scr-family (prototype: secretin receptor), and the family 3 or mGluR family (prototype: metabotropic glutamate receptors.

Family 1 is divided according to the size and chemical nature of the corresponding agonists, as well as the mode of ligand binding. Family 1a accommodates the β-adrenoceptor-type receptors that are activated by small ligands, such as biogenic monoamines, opiates, nucleotides and small peptides that bind to a transmembrane cavity formed by helices 3, 4, 5 and 6. Family 1b is composed of receptors stimulated by oligopeptides and proteins, such as IL-8 (interleukin-8), cytokines and thrombin. The ligand binding epitope is located in the extracellular loop region. Family 1c receptors recognise glycoprotien hormones, such as LH (luteinising

hormone), TSH (thyroid-stimulating hormone), and FSH (follicle-stimulating hormone), while their ligand binding site is centered in a large extracellular N-terminal domain.

Family 2 receptors are distinct from the rho-family receptors in that they bind large molecules like glucagons, secretin, PTH (parathyroid hormone), VIP (vasointestinal peptide), or CRF (corticotropin-releasing factor). Comparable to family 1c receptors, the secretin family utilises a large N-terminal domain for ligand binding. Family 3 receptors are unique in that they possess a large extracellular N-terminal domain of several hundred residues that constitutes the binding site for smallish ligands, such as a single divalent Ca^{2+} cation, glutamate, GABA (γ-amino butyric acid) and pheromones.

Examples of GPCR members and GPCR subfamilies are given in the following Table 1:

Table 1

GPCR	code	Native ligand (peptide/protein)	Nature of the ligand
angiotensin receptors	AT_1, AT_2	angiotensin II (AII)	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe
bombesin receptors	BB1 - BB4	bombesin, neuromedin B, gastrin-releasing peptide	14 aa peptide amide
bradykinin receptors	B_1, B_2	bradykinin (BK)	Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg
C3a receptor	C3aR	C3a anaphylatoxin	protein
C5a receptor	C5aR	C5a anaphylatoxin	protein
CC chemokine receptors	CCR1 - CCR9	chemokines	proteins
CXC chemokine receptors	CXCR1 - CXCR5	chemokines	proteins
cholecystokinin/gastrin receptors	CCK_A, CCK_B	cholecystokinin (CCK), gastrin	33 aa peptide amide, 17 aa peptide amide
endothelin receptors	ET_A-ET_B	endothelin-1 (ET-1), ET-2, ET-3	21 aa peptides
alpha factor pheromone receptor	STE2, STE3	fungal mating pheromones	13 aa peptide
fMet-Leu-Phe receptor	fMLP-R	Formylpeptide (fMLP)	FMet-Leu-Phe

GPCR	code	Native ligand (peptide/protein)	Nature of the ligand
galanin	GAL1, gal2, gal3	galanin	30 aa peptide
melanocortin receptors & ACTH receptor	MC_1,MC_3,MC_4,MC_5 MC_2 = ACTH receptor	melanocortin (MSH) adrenocorticotropic hormone (ACTH), corticotropin	39 aa peptide
neuropeptide Y receptor	Y_1 - Y_6	neuropeptide Y (NPY), peptide YY (PYY), pancreatic polypeptide (PP)	36 aa peptide amide (NPY)
neurotensin receptor	NTS1, nts2	neurotensin	13 aa peptide
opioid receptors	δ	[Met]-enkephalin, [Leu]-enkephalin	Tyr-Gly-Gly-Phe-Met/Leu
	κ	dynorphin A	17 aa peptide
	μ	β-endorphin, Lipotropin C fragment	31 aa peptide
nociceptin receptor	ORL1	nociceptin, orphanin FQ	17 aa peptide
somatostatin receptors	sst_1 - sst_5	somatostatin	cyclic 14 aa peptide
tachykinin receptors	NK_1	substance P	11 aa peptide
	NK_2	neurokinin A (NKA), substance K, neuromedin L	His-Lys-Thr-Asp-Ser-Phe-Val-Gly-Leu-Met-NH_2
	NK_3	neurokinin B (NKB), neuromedin K	Asp-Met-His-Asp-Phe-Phe-Val-Gly-Leu-Met-NH_2
thrombin/protease-activated receptors	PAR1, PAR2, PAR3, PAR4	thrombin, trypsin, factor Xa	protein
vasopressin receptors	V_{1A}, V_{1B}, V_2	vasopressin	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH_2
oxytocin receptor	OT	oxytocin	Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly-NH_2
vasotocin receptor	VT	vasotocin	Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Arg-Gly-NH_2

GPCR	code	Native ligand (peptide/protein)	Nature of the ligand
orexin receptors	OX_1, OX_2	orexin A/B	33 aa/28 aa peptide amides
FSH receptor	FSH receptor	follicle-stimulating hormone (FSH)	protein
LSH receptor	LSH receptor	lutropin, choriogonadotropic hormone, luteinizing hormone	protein
TSH receptors	TSH receptor	thyrotropin, thyroid-stimulating hormone	protein
LHRH receptor	LHRH receptor	gonadotropin-releasing hormone (GnRH), luteinizing hormone-releasing hormone (LHRH)	pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2
thyrotropin-releasing hormone & secretagogue receptors	TRH_1, trh_2	thyrotropin-releasing hormone/factor (TRH/F)	pGlu-His-Pro-NH_2
GHS receptor	$GHSR_{1a}$, $GHSR_{1b}$	growth hormone secretagogues (GHS)	oligopeptides
calcitonin/calcitonin gene-related peptide receptors	CGRPR	calcitonin, calcitonin gene-related peptide (CGRP)	32 aa peptide amide
amylin receptor	amylin receptor	amylin	37 aa peptide amide
adrenomedullin receptor	adrenomedullin receptor	adrenomedullin	52 aa peptide amide
corticotropin-releasing factor receptor	CRF_1, CRF_2	corticotropin-releasing factor (CRF)	41 aa peptide amide
gastric inhibitory peptide receptor	gip receptor	gastric inhibitory peptide (GIP)	42 aa peptide
glucagon/glucagon-like peptide receptor	GLP1	glucagon	29 aa peptide

GPCR	code	Native ligand (peptide/protein)	Nature of th ligand
growth-hormone-releasing hormone receptor	GHRH receptor	growth-homone-releasing homone/factor (GHRH/GRF)	44 aa peptide amide
parathyroid hormone receptor	type 1, type 2	parathyroid hormone (PTH)	84 aa peptide
secretin receptor	secretin receptor	secretin	27 aa peptide amide
vasoactive intestinal peptide & PACAP receptor	$VPAC_1$, $VPAC_2$, PAC_1	vasoactive intestinal peptide (VIP) pituitary adenylate cyclase activating peptide (PACAP)	28 aa peptide amide 38 aa peptide

GPCR arrays

A GPCR array according to the invention has a multiplicity of individual GPCR polynucleotide spots, stably associated with a surface of a solid support. Each spot on the GPCR array comprises a GPCR polynucleotide composition, wherein the polynucleotide regions within the composition are of known identity, usually of known sequence, as described later on in detail. The GPCR polynucleotide spots may be of convenient shape but most often circular, oval or any other suitable shape. The GPCR polynucleotide spots may be arranged in any convenient pattern across the surface of the solid support, such as in row or columns to form a grid, in a circular pattern and the like. Preferably the pattern of GPCR polynucleotide spots is arranged as a grid to facilitate the evaluation of the results obtained from the analyses in which the GPCR array is used.

The GPCR array according to the invention may be of a flexible or rigid solid support and the GPCR polynucleotide spots are stably associated thereto. By stably associated is meant that the GPCR polynucleotide spots will be associated in their position on the solid support during the analysis in which the GPCR array is used, such as during different hybridisation, washing and detection conditions. The GPCR polynucleotide regions contained in the spots may be covalently or non-covalently associated to the surface of the solid support. Methods how to covalently or non-covalently bind the GPCR polynucleotide regions to the surface of the solid support are well known for a person skilled in the art and may be found in e.g. *Ausubel et al.*, Current Protocols in Molecular Biology, Greene Publishing Co. NY, (1995).

The solid support to which the individual GPCR polynucleotide spots are stably associated to is made of a flexible or rigid material. By flexible is meant that the support is capable of being bent or folded without breakage. By rigid is meant that the support is solid and does not readily bend, i.e. the support is not flexible. The support may be fabricated from a variety of materials, including plastics, ceramics, metals, gels, nitrocellulose, nylon, glass and the like.

The array may be produced according to any convenient methodology, such as preparing or obtaining the polynucleotides and then stably associate them with the surface of the support or growing them directly on the support. A number of array configurations and methods for their production are known to those skilled in the art and disclosed in e.g. US 5,445,934, US 5,532,128, US 5,556,752, US 5,242,974, US 5,384,261, US 5,405,783, US 5,412,087, US 5,424,186, US 5,429,807, US 5,436,327, US 5,472,672, US 5,527,681, US 5,529,756, US 5,545,531, US 5,554,501, US 5,561,071, US 5,571,639, US 5,593,839, US 5,599,695, US 5,624,711, US 5,658,734 and US 5,700,637.

The solid support of the invention may have several configurations ranging from a simple to a more complex configuration depending on the intended use of the GPCR array. The size and thickness of the GPCR array is not critical as long as the GPCR array will function in the expected way and as long as the results obtained after use of the GPCR array are not changed. The number and amount of the GPCR polynucleotide spots is dependent on the intended use of the GPCR arrays as well as the detection system use to determine the expression profile of the biological material being evaluated by the aid of the GPCR array. The number of the GPCR polynucleotide spots may vary from about 2 to about 100,000 such as e.g. from about 2 to about 50,000, from about 10 to about 25,000, from about 100 to about 10,000, from about 100 to about 5,000, from about 100 to about 1,000, from about 400 to about 600 or about 500 GPCR polynucleotide spots, or at least 2 such as, e.g. at least 10, at least 25, at least 50, at least 100, at least 300, at least 400, at least 500 or at least 600 spots, or even more than 100,000 spots. The limitations of the number of the GPCR polynucleotide spots are dependent on the way in which the evaluation of the expression profile of the biological material is performed. The amount of the GPCR polynucleotide regions present in the GPCR polynucleotide spot may vary and the amount will be sufficient to provide adequate hybridisation and detection of the target nucleic acid. Generally the GPCR polynucleotides will be present in each spot at a concentration corresponding to an amount of 1 ng – 10 µg per µl or less than 1 µg of the polynucleotide. Normally, only 1 GPCR polynucleotide region is present in each spot.

The copy number of the GPCR polynucleotide present in each GPCR polynucleotide spot will be sufficient to provide enough hybridisation for a target nucleic acid to yield a detectable signal, and generally range from about 50 fmol or less.

An important feature of the GPCR array is i) that the majority of the GPCR polynucleotide spots represent GPCR families which have GPCR members, ii) the GPCR polynucleotide regions present in the GPCR polynucleotide spots are made up from non-conserved regions of the GPCR family members, and iii) at least two GPCR polynucleotide regions representing each GPCR family member are present on the GPCR array. The two or more GPCR polynucleotide regions from one GPCR family member are chosen in such a way that they are non-overlapping regions. The use of two or more GPCR polynucleotide regions on the GPCR array ensures a proper expression profile from the same GPCR polynucleotide. In general, prior art arrays have suffered from the technical problem that they are not fully reliable in the sense that they produce a certain level of both false negative and false positive results. This technical problem has been solved with the present invention by the use of at least two regions of the same GPCR polynucleotide family member. Such array design greatly increases the reliability of the results generated by the array, as each determination has at least one double check or verification (positive and negative control). Also, this double check may be effected in a controlled manner and to a pre-determined level by selection of an appropriate number of regions of a GPCR polynucleotide family member. The level of double check, i.e. the number of regions, may be selected specifically for the intended use and depends on several factors, which includes, but are not limited to, the length of the polynucleotide regions spotted, the degree of intra-family identity, the degree of interspecies identity, the array hybridisation conditions, the characteristics of the biological polynucleotide test sample etc.

The number of non-conserved polypeptide regions to be chosen is dependent i.a. on the length of the corresponding GPCR polynucleotide. Additionally, the ability to identify expression of GPCR polynucleotide which for some reason have a mutation or a deletion may increase by the use of more than one non-conserved polynucleotide region from each member of a GPCR family. Generally the non-conserved regions will be found in the three extracellular loops, the three intracellular loops or in the N- or C-terminal end of the GPCR.

Preferably, each type of spot as defined by its content of polynucleotide region is present in a number of copies, such as 2, 3, 4, 5 or 6, in order to enhance the reliability of the results obtained in the use of the array. When multiple spots of the same type are used, the mean value of the results obtained is calculated and used.

The array according to the invention has several different regions of a GPCR polynucleotide family member, which may be polynucleotide regions from the same polynucleotide strand and the regions differ at least by one nucleotide.

The non-conserved regions corresponding to a specific GPCR member of a GPCR family are preferably selected such that the selected GPCR regions have the ability to hybridise to the corresponding polynucleotide from more than one species. One selected GPCR region may be used for the identification of the expression profile of a certain GPCR polynucleotide in several biological materials obtained from several species such as e.g. humans, mice and rats. By this feature the functionality of the GPCR array increases such that solely one type of GPCR array is needed for the evaluation of the expression profiles of GPCR polynucleotide in biological materials obtained from several species. One GPCR polynucleotide spot will hybridise specifically to one single member of a GPCR family due to the selected non-conserved region of that particular GPCR. By the use of such a strategy for the development of the GPCR array, several different biological materials such as, e.g. material obtained from different species of animals may be used and compared for their expression profile of GPCR polynucleotides using only one type of GPCR array. The same strategy also applies for plants, fungi, microorganisms etc.

Other polynucleotide spots (control spots), which may be present on the GPCR array, include spots comprising genomic DNA, housekeeping genes, negative and positive control polynucleotides and the like. These polynucleotide spots comprise polynucleotides, which are not unique, i.e. they are not polynucleotide regions corresponding to GPCR polynucleotides. They are used for calibration or as control polynucleotides, and the function of these polynucleotide spots are not to give information of the expression of these polynucleotides, but rather to provide useful information, such as background or basal level of expression to verify that the analysis and the expression profiles obtained are relevant or not. Furthermore these control spots may serve as orientation spots.

The GPCR polynucleotides of interest in the present context are those, which encode seven transmembrane polypeptides involved in transducing a signal across a biological membrane. Examples of suitable GPCR polynucleotide families are adrenergic, adenosine, dopaminergic, histaminergic, opioid and serotonin GPCR's and any other seven transmembrane polynucleotides capable of transducing a signal across biological membranes.

The GPCR polynucleotides to be stably associated to the solid support may be of DNA, RNA, cDNA, natural, synthetic, semi-synthetic origin or chemical analogous such as LNA or PNA. The GPCR polynucleotides may be obtained from one or more biological material such as an organism, an organ, a tissue and/or a cell

and/or produced by a cell culture. The biological material may be obtained from any kind of organism, such as a microorganism, a plant, a fungus (e.g. yeast, mushrooms), animal or tissue. Examples of animals from which one or more biological material may be obtained are humans, rats, mice, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, apes, monkeys, cats and sheep.

The GPCR polypeptides involved in transducing a signal across biological membranes may be located e.g. in an organ such as heart, liver, prostate, brain, kidney, lung etc., tissue such as nerve, muscle, connective, etc., and/or they may be found in the cells such as e.g. in the nucleus, endoplasmatic reticulum, Golgi complex, endosome, lysosome, peroxisome, mitochondria, cytoplasm, plasma membrane, cytoskeleton.

The length of the GPCR polynucleotides present in the GPCR polynucleotide spot is selected in such a manner that the length is sufficient to provide a strong, specific and reproducible signal. The length will typically vary from about 3 to about 9,000 nucleotides such as e.g. from about 3 to about 6,000, from about 3 to about 3,000, from about 10 to about 1,500, from about 50 to about 1,000, from about 100 to about 800, from about 200 to 750, from about 200 to 700, from about 200 to 500, from about 250 to 400 or preferably from about 300 to 350. However, the length of the GPCR polynucleotides present on the GPCR array is shorter than the length of the mRNA to which it corresponds. As such, the GPCR polynucleotide represents a part of the full-length cDNA to which it corresponds. The length of the GPCR polynucleotide region present in the GPCR spot is dependent on the number of polynucleotides in the selected GPCR family member.

The non-conserved regions of GPCR polynucleotide regions contained in a GPCR polynucleotide composition may be single or double stranded non-conserved polynucleotide regions.

The GPCR polynucleotide composition also comprises an excipient. Suitable excipients are solvents like e.g. water or any other aqueous medium, pH adjusting agents like buffering agents, stabilising agents, hybridising agents, colouring agents, labelling agents and the like. In general, the excipients used are inert, i.e. they do not have any polynucleotide related effect.

Method of selecting the GPCR polynucleotides

According to the present invention an important feature of the GPCR array is that the majority of the GPCR polynucleotide spots are made up from family members of GPCR polynucleotides and the GPCR polynucleotide regions present in the GPCR polynucleotide spots are made up from non-conserved regions of the GPCR polynucleotides. The sequences of the GPCR families may be found in GenBank (http.www.ncbi.nih.gov) and downloaded prior to sequence comparison. The

sequence comparisons may be performed using any of the methods mentioned above. An example is given in Example 1 herein.

Additionally, the GPCR array preferably represents at least two different GPCR polynucleotide members of a family and/or at least two different GPCR families.

Furthermore, at least two of the GPCR polynucleotide spots are made up from one GPCR polynucleotide of one GPCR polynucleotide family member. The GPCR polynucleotide regions present in the two GPCR spots are made up from regions of one and the same GPCR polynucleotide and the regions are at least non-overlapping with a distance of at least one nucleotide from each other. The GPCR polynucleotide regions are selected in such a way that they are non-conserved regions within the same GPCR family member (intrafamily) and the regions have at least 50% identity between different species (interspecies). The strategy how to find and identify potential GPCR regions useful to stably associate onto the surface of the GPCR array will be described in detail hereinafter.

In one embodiment of the invention relates to a GPCR array comprising a multiplicity of individual GPCR polynucleotide spots stably associated with a surface of a solid support, wherein an individual GPCR polynucleotide spot comprises a GPCR polynucleotide composition, and the spots represent at least two different regions of a GPCR polynucleotide family member. The GPCR polynucleotide composition comprises a non-conserved region of a GPCR polynucleotide family member. The non-conserved regions of a GPCR polynucleotide family member is a stretch of nucleotides with an average length of from about 3 to about 9,000 nucleotides such as, e.g. from about 3 to about 6,000, from about 3 to about 3,000, from about 5 to 1,500, from about 10 to about 1,000, from about 50 to about 1,000, from about 100 to about 1,000, from about 200 to 750, from about 200 to 700, from about 200 to 500, from about 250 to 400 or from about 300 to 350. The non-conserved regions of a GPCR polynucleotide family member is a region of nucleotides which has less than 90% such as, e.g. less than 85%, less than about 80%, less than about 75%, less than about 70%, less than about 65%, less than about 60%, less than about 55% or less than about 50% intrafamily identity, which means less than 90% (or, alternatively 85%, 80%, 75% or 70%) identity between polynucleotides classified as member of a specific GPCR family, see example 1. The homology between members of a certain GPCR family may be determined using the methods mentioned above.

The two or more different non-conserved polypeptide regions corresponding to one GPCR member may be identified using the same strategy and they may at least be non-overlapping regions as mentioned above.

The non-overlapping regions may be selected from just one non-conserved region in case the polynucleotide GPCR family member contains just one non-conserved region.

The non-conserved polypeptide regions may furthermore be selected on the basis of homology of specific regions between different species (interspecies), such as between species of microorganisms, fungi, plants or animals such as e.g. humans, rats, mice, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, apes, monkeys, cats and sheep.

The non-conserved region of the GPCR family member may have at least 50% interspecies identity such as, e.g. at least about 60%, at least about 65%, at least about 70%, at least about 75% or at least 80% interspecies identity. By the use of such as strategy in which the non-conserved regions are selected on the basis of their interspecies identity, it is possible to use the same type of GPCR array for detection of GPCR expression profiles in several different species as long as the selected non-conserved regions shares such a high degree of homology to enable hybridisation.

The difference between the percent interspecies identity and the percent intrafamily identity (percent interspecies identity minus percent intrafamily) should be a least 5%, at least 10%, at least 15%, at least 20% and at least 25%. By such a strategy it should be possible to select hybridisation conditions, which may preferentially bind interspecies related polynucleotides and avoid the binding of intrafamily related polynucleotides.

The non-conserved region of the GPCR polynucleotide used according to the invention will generally be a single stranded polynucleotide and shorter than the mRNA to which it corresponds.

In one embodiment of the array of the invention, at least one non-conserved polynucleotide region is present in the form of sense single-strands in a spot. In conventional arrays the polynucleotides are usually present in double-stranded form, which is denatured by heating prior to contacting with the biological sample to make the sense strand available for binding with the sample polynucleotides. It is believed that in such conventional arrays a certain variable and non-controlled level of the double-stranded polynucleotides on the array does not in fact separate sufficiently to allow hybridisation with the sample polynucleotide. In comparison, an array wherein the sense polynucleotide is present in a single-stranded form has the advantage that all strands are available for hybridisation thus resulting in an increased and more reproducible level of binding and array response. Preferably, the said non-conserved polynucleotide region present in the form of sense single-strands in a spot is also present in the form of antisense single-strands in a separate spot. This embodiment involves the advantage that the spot containing the antisense

strands serves as a negative control for a positive determination in the spot containing the corresponding sense strands. Thus, this embodiment of the invention significantly increases the reliability of the results obtained in the use of the array. Furthermore, when only the sense sequence of the polynucleotide region is included in the array, it is necessary to identify the sense strand of the double-stranded polynucleotide region during the preparation of the array. In comparison, when both the sense and the anti-sense strand of the polynucleotide region are included as separate spots, identification of which is which is avoided hence facilitating the preparation of the array, which then can be carried out using standard methods.

When the array is to be used for screening biological polynucleotide samples originating exclusively or primarily from one animal species, the GPCR polynucleotide regions present in the spots preferably originate from the same species. This is preferred because it will give an optimum level of identity between the polynucleotides of the spot on the one side and the polynucleotides of the biological material on the other side, and hence a more reliable determination.

Method of preparing a GPCR array

The GPCR array may be prepared (produced) using any convenient method and several methods are well known for a person skilled in the art, such as standard procedures according to e.g. *Sambrook et al.*, Molecular cloning: A laboratory manual 2nd edition, Cold Spring Harbour Laboratory Press, New York.

One means of preparing the GPCR array is i) synthesising or otherwise obtaining the above mentioned non-conserved GPCR polynucleotide regions, ii) preparing the GPCR polynucleotide compositions to be used in each spot and then iii) depositing in the form of spots the polynucleotide compositions comprising the non-conserved GPCR polynucleotide regions onto the surface of the solid support, see also Examples 2-5. The non-conserved GPCR polynucleotide regions may be of DNA, RNA, cDNA, natural, synthetic, semi-synthetic origin or chemical analogous such as LNA or PNA. The non-conserved regions may be obtained from any biological material such as e.g. tissues or cells and/or produced by a cell culture. The biological material may be an organism, such as a microorganism, plant, fungus (e.g. yeast or mushrooms) or animal. If the organism is an animal it may be selected from a group consisting of humans, rats, mice, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits or sheep.

The non-conserved GPCR polynucleotide regions may be prepared using any conventional methodology such as automated solid phase synthesis protocols, PCR using one or more primers specific for the non-conserved GPCR polynucleotide regions and the like. In general, PCR is advantageous in view of the large numbers of

non-conserved GPCR polynucleotide regions that must be generated for each GPCR array. The amplified non-conserved GPCR polynucleotide regions may further be cloned in any suitable plasmid vector to enable multiplication and storage of the amplified non-conserved GPCR polynucleotide regions (see Examples 3-4).

The prepared non-conserved GPCR polynucleotide regions may be spotted onto the solid support using any convenient methodology, including manual and automated techniques, e.g. by micro-pipette, ink jet pins etc. and any other suitable automated systems. An example of an automated system is the automated spotting device Beckman Biomek 2000 (Beckman Instruments, USA). The ready GPCR arrays may then be stored at suitable conditions until use.

Method for the determination of GPCR expression profiles

Determination of GPCR expression profiles typically means determination of the expression level of multiple mRNA's, all of them corresponding to GPCR polynucleotides. The detection limit of the expression level of mRNA may be approximately 0.2 ng or less of total RNA of the biological material used to hybridise each individual GPCR polynucleotide spot.

The expression profiles can be produced by any means known in the art, including but not limited to the methods disclosed by e.g. *Liang et al.*, Science 1992 **257** 967-971; *Ivanova et al.*, Nucleic Acids Res. 1995 **23** 2954-2958; *Guilfoyl et al.*, Nucleic Acids Res. 1997 **25** (9) 1854-1858; *Chee et al.*, Science 1996 **274** 610-614; *Velculescu et al.*, Science 1995 **270** 484-487; *Fiscker et al.*, Proc. Natl. Acad. Sci. USA 1995 **92** (12) 5331-5335; and *Kato*, Nucleic Acids Res. 1995 **23** (18) 3685-3690.

The hybridisation conditions under which the biological polynucleotide sample is contacted with the array of the invention may vary and are selected to suit the characteristics of the specific array / sample system as well as the purpose of the use of the array. The hybridisation conditions selected depend e.g. on the species from which the biological sample originates, the length of the polynucleotide regions in the spots, the number of polynucleotide regions from the same GPCR family member, the level of intra-family identity, the level of interspecies identity, the array reaction conditions, such as the type of solid support used, the type of system used for linking the GPCR polynucleotides to the solid support and the type of hybridisation chamber used; the characteristics of the biological polynucleotide test sample, such as purity, concentration, expected amount of cDNA, the quality of the cDNA etc. Depending on the before-mentioned factors, the hybridisation conditions may be adjusted to each individual array system. Depending on the said factors it is in general possible to use low stringent, medium stringent and high stringent hybridisation conditions. Preferably, high stringent conditions are used for human samples and medium stringent samples

are used for rat and mouse samples. In connection with the present invention, an example of low stringent conditions is 40% formamide 1 M Na and a temperature of 37°C. An example of medium stringent conditions is 1 M Na and a temperature of 55°C. An example of high stringent conditions is 1 M Na and a temperature of 65°C. For all types of hybridisation, the incubation period is preferably more than 16 hours, more preferably more than 20 hours, and most preferably more than 24 hours.

According to one embodiment of the invention the GPCR array will be used for the evaluation of the expression profile of one or more biological materials or a mixture of biological materials. The method for the determination of a GPCR polynucleotide expression profile in a biological material or in a mixture of biological materials comprises obtaining a polynucleotide from the biological material(s), labelling said polynucleotide to obtain a labelled target polynucleotide sample, contacting at least one labelled target polynucleotide sample with an array as defined above under conditions which are sufficient to produce a hybridisation pattern and detecting said hybridisation pattern to obtain the GPCR polynucleotide expression profile of the biological material or the mixture of biological materials. The GPCR expression profile in the biological material can thus be determined to correspond to the expression of e.g. GPCR's such as adrenergic, adenosine, dopaminergic, histaminergic, opioid and serotonin GPCR's members, or any other polynucleotides encoding polypeptides capable of transducing a signal across biological membranes. The biological material or the mixture of biological materials may be in a non-stressed or a stressed stage. The stress may directly or indirectly influence the GPCR expression profile and thereby the polynucleotides identified which react upon that type of stress. The stress may be caused by a disease or a condition such as e.g. Asthma, cystic fibrosis, chronic obstructive pulmonary disease and rhinorrhea, convulsions, vascular spasms, coronary artery spasms, renal disorders, polycystic kidney disease, bladder spasms, urinary incontinence, bladder outflow obstruction, irritable bowel syndrome, gastrointestinal dysfunction, secretory diarrhoea, ischaemia, cerebral ischaemia, ischaemic hearth disease, angina pectoris, coronary hearth disease, traumatic brain injury, psychosis, anxiety, depression, dementia, memory and attention deficits, drug addiction and/or abuse, including cocaine or tobacco abuse, Parkinson's disease, Alzheimer's disease, dysmenorrhoea, narcolepsy, Reynaud's disease, intermittent claudication, Sjorgren's syndrome, migraine, arrhythmia, hypertension, absence seizures, myotonic muscle dystrophia, xerostomi, diabetes type II, hyperinsulinemia, premature labour, baldness, cancer, schizophrenia or psychosis.

A variety of disorders associated with the neural system, for example eating disorders, obsessive-compulsive disorders, panic disorders, alcoholism, pain, memory

deficits and anxiety. Included among these disorders are disorders such as pseudodementia or Ganser's syndrome, migraine pain, bulimia, obesity, pre-menstrual syndrome or late luteal phase syndrome, post-traumatic syndrome, memory loss, memory dysfunction, social phobia, attention deficit hyperactivity disorder, chronic fatigue syndrome, premature ejaculation, erectile difficulty, anorexia nervosa, disorders of sleep, autism, mutism, trichotillomania or mood syndrome.

Auto-immune diseases, e.g. Addison's disease, alopecia areata, Ankylosing spondylitis, haemolytic anaemia (anaemia haemolytica), pernicious anaemia (anaemia perniciosa), aphthae, aphthous stomatitis, arthritis, arteriosclerotic disorders, osteoarthritis, rheumatoid arthritis, aspermiogenese, asthma bronchiale, auto-immune asthma, auto-immune haemolysis, Becket's disease, Boeck's disease, inflammatory bowel disease, Burkett's lymphoma, Chron's disease, chorioiditis, colitis ulcerosa, Coeliac disease, cryoglobulinemia, dermatitis herpetiformis, dermatomyositis, insulin-dependent type I diabetes, juvenile diabetes, idiopathic diabetes insipidus, insulin-dependent diabetes mellisis, auto-immune demyelinating diseases, Dupuytren's contracture, encephalomyelitis, encephalomyelitis allergica, endophthalmia phacoanaphylactica, enteritis allergica, auto-immune enteropathy syndrome, erythema nodosum leprosum, idiopathic facial paralysis, chronic fatigue syndrome, febris rheumatica, glomerulo nephritis, Goodpasture's syndrome, Graves' disease, Hamman-Rich's disease, Hashimoto's disease, Hashimoto's thyroiditis, sudden hearing loss, sensoneural hearing loss, hepatitis chronica, Hodgkin's disease, haemoglobinuria paroxysmatica, hypogonadism, ileitis regionalis, iritis, leucopoenia, leukaemia, lupus erythematosus disseminatus, systemic lupus erythematosus, cutaneous lupus erythematosus, lymphogranuloma malignum, mononucleosis infectiosa, myasthenia gravis, traverse myelitis, primary idiopathic myxedema, nephrosis, ophthalmia symphatica, orchitis granulomatosa, pancreatitis, pemphigus, pemphigus vulgaris, polyarteritis nodosa, polyarthritis chronica primaria, polymyositis, polyradiculitis acuta, psoreasis, purpura, pyoderma gangrenosum, Quervain's thyreoiditis, Reiter's syndrome, sarcoidosis, ataxic sclerosis, progressive systemic sclerosis, scleritis, sclerodermia, multiple sclerosis, sclerosis disseminata, acquired spenic atrophy, infertility due to antispermatozoan antibodies, thrombocytopenia, idiopathic thrombocytopenia purpura, thymoma, acute anterior uveitis, vitiligo, AIDS, HIV, SCID and Epstein Barr virus associated diseases such as Sjorgren's syndrome, virus (AIDS or EBV) associated B cell lymphoma, parasitic diseases such as Lesihmania, and immune-suppressed disease states such as viral infections following allograft transplantations, graft vs. Host syndrome, transplant rejection, or AIDS, cancers, chronic active hepatitis diabetes, toxic chock syndrome, food poisoning, and transplant rejection.

These are examples and are not intended to limit the invention in any way.

The analysis of the expression profile includes several steps of procedures in which well known techniques are used, such as those mentioned in e.g. *Sambrook et al.*, Molecular Cloning: A Laboratory approach, Cold Spring Harbour Press, NY (1987), and in *Ausubel et al.*, Current Protocols in Molecular Biology, Greene Publishing Co. NY, (1995).

The biological material to be evaluated needs to be identified and isolated such as e.g. described in Example 2. For the ability to perform the analysis cDNA are generally produced from isolated total RNA or polyA RNA (mRNA). The total RNA/mRNA can be isolated using a variety of techniques. Numerous techniques are well known (see e.g. *Sambrook et al.* and *Ausubel et al.*, op cit.). In general, these techniques include a first step of lysing the cells and then a second step of enriching for or purifying RNA.

The isolated total RNA/mRNA are reversed transcribed using a RNA-directed DNA polymerase, such as "reverse transcriptase" isolated from such retroviruses as AMV, MoMuLV or recombinantly produced. Many commercial sources are available (from e.g. Perkin Elmer, New England Biolabs, Stratagene Cloning Systems).

Preferably the mRNA is reversed transcribed into cDNA and at the same time a label is incorporated for later detection of the hybridised amplified products on the GPCR array. The amplification by PCR may be performed according to Example 2. The label may vary dependent on the system to be used for the detection and several labels are well known in the area of molecular biology (e.g. radioactive labels, fluorescent labels, colouring labels, chemical labels etc.)

The labelled cDNA is then denaturised and used for hybridisation on the GPCR array. The hybridisation conditions vary and are dependent on the aim with the expression profile obtained after the hybridisation. One example is found in Example 6. After hybridisation of the labelled cDNA, the GPCR array is washed to remove the cDNA, which have not hybridised to the GPCR and the hybridised labelled cDNA are detected by a suitable means and an expression profile obtained.

In a second embodiment of the present invention two GPCR arrays will be used for the evaluation of the expression profiles in at least a first and a second biological material. The expression profiles of the first and the second biological material are compared to each other to identify any differences between the first and the second expression profile. The analysis comprising obtaining a first GPCR expression profile of the first biological material as described above, obtaining a second GPCR expression profile of the second biological material as described for the first biological material, comparing the first and the second GPCR expression profile to

identify any differences in the GPCR expression profiles between the first and the second GPCR expression profile. The first and the second biological material may be of the same origin of different origins, for example two livers from the same animal species or two lungs from the same animal or from two animals of the same species etc. In one embodiment of the invention, the first and the second biological material are in two different stages, i.e. the first biological material is non-stressed and the second biological material is stressed. The second biological material may be stressed in such as way that at least a different GPCR expression profile will be obtained. The stress may directly or indirectly influence the GPCR expression profile.

The GPCR's to be influenced by the stress of the second biological material may be any GPCR belonging to any of the overall group of Rhodopsin-like receptors, secretin-like, metabotropic glutamate/pheromone, Frizzled/smoothened and a large group of unclassified GPCR's. The groups of Rhodopsin-like receptors are typically divided into smaller subgroups like e. g. adrenergic, adenosinergic, dopaminergic, histaminergic, opioid, serotonergic and peptide hormone receptor like e.g. angiotensin, bradykinin, chemokine, endothelin and melanocortin, neuropeptide Y, somatostatin, tachykinin, galanin, orexin, rhodopsin, olfactory, prostaglandin, nucleotide-like and purinoceptor. Some of the GPCR belonging to the group of secretin-like receptorsion includes calcitonin, corticotropin releasing factor, gastric inhibitory peptide, glucagons, growth hormone-releasing hormone, parathyroid hormone, PACAP, secretin, vasoactive intestinal polypeptide, diuretic hormone, EMR1, latrotoxin. The metabotropic glutamate/pheromone GPCR's are exemplified by the metabotropic glutamate, extracellular calcium sensing, GABA-B and putative pheromone receptors.

Furthermore, the GPCR expression profile of the second biological material may be directly or indirectly related to a disease, chemical pre-treatment, environmental influences or other physiological or pathophysiological changes in the biological material. The chemical treatment may be selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities.

Examples of diseases or conditions that might influence the GPCR expression profile of the second biological material are those mentioned above. As an example the use of *in vivo* models such as e.g. a rat model in which at least a first and a second experimental group are used. The first group is non-stressed and the second group stressed in such a way that the expression of one or more GPCR polynucleotides are influenced in such as way that an increase or a decrease of the expression is obtained, when the expression profiles are analysed using the GPCR array and the method according to the invention. The second group may be either permanently stressed or stressed during a certain period of time and after the period

of stress one or more biological materials obtained from the second group and the GPCR expression profile determined

In another embodiment of the present invention, the GPCR array is used for the evaluation of the expression profiles in at least a first and a second biological material, each material being labelled with a unique label (e.g. Cy3 and Cy5 for each sample, respectively). The procedure is as described above.

In a third embodiment according to the invention, the GPCR array will be used for the identification of a therapeutic, prophylactic or toxic agent involved in the response of GPCR polypeptides in a biological material, said method comprises obtaining a first GPCR expression profile of a first biological material as described above, obtaining a second GPCR expression profile of a second biological material as described above, treating the first and/or the second biological material with a test compound; obtaining a third and/or a fourth GPCR expression profile of the treated second biological material as described above, comparing the first, second, third and/or fourth GPCR expression profiles, and identifying any difference in the GPCR expression profile so as to identify any therapeutic, prophylactic or toxic response of the test compound on the GPCR polynucleotide. The first biological material may typically be a material in a healthy or normal condition whereas the second biological material typically may be in a diseased or not normal state. The first and the second biological material may have the same kind of origin.

The first biological material may be in a non-stressed state and the second biological material may be in a stressed state and the stress may directly or indirectly influence the GPCR expression profile between the first and the second biological materials. The GPCR polynucleotide family which is influenced by the stress is selected from the group consisting of voltage-gated GPCR's, Gap-junction GPCR's, ligand-gated GPCR's, heat-activated GPCR's, intracellular GPCR's, GPCR gated by intracellular ligands such as cyclic nucleotide-gated channels or calcium-activated GPCR's, and the GPCR expression profile of the second biological material is directly or indirectly related to a disease, a chemical or biological pre-treatment, environmental influences or other physiological or pathophysiological changes.

The disease may be anyone of those mentioned above.

The test compound may be a chemical or a biological compound including therapeutic, prophylactic and/or toxic chemical entities, physiologically chemical entities, substances affecting a biological function, hormones, vitamins, nutrients, pesticides, fungicides, bacteriocides and the like. The method according to the third embodiment of the invention is used to identify potential therapeutic, prophylactic and/or toxic agents useful for the treatment of diseases caused by an alteration in the expression profile of the GPCR polypeptides. One example is the use of a biological

model, such as a rat model in which a first, second, third and/or fourth group are used. The first and third group is non-stressed and the second and fourth group stressed in such a way that the expression of one or more GPCR polynucleotides are influenced in such as way that an increase or a decrease of the expression is obtained. The third and fourth groups are treated with a test compound.

In a fourth embodiment the invention will be used in diagnostic methods to enable the determination in differences of GPCR expression profiles between two different biological material, said method comprises obtaining a first GPCR expression profile of a first biological material as described above, obtaining a second GPCR expression profile of a second biological material as described above, comparing the first and second GPCR expression profiles, and identifying any difference in the GPCR expression profile. Preferably a disease, such as chronic pain, myotonia, multiple sclerosis, and rheumatoid arthritis, directly or indirectly influences the difference between the GPCR expression profile of the first and the second GPCR expression profile. Furthermore, the GPCR expression profile from more than two different biological materials are compared, such as biological materials, which are in different stages of a disease. The diagnostic method may be useful in the determination of diseases directly or indirectly caused by different GPCR expression profiles and by the use of such a method there will be an enhanced possibility to start the treatment of the disease at an early stage of the disease.

GPCR kits

The invention also relates to kits comprising the above-mentioned GPCR array.

The kit may be used according to the above-mentioned methods for the determination of GPCR expression profiles in a biological material as defined above.

The kit comprises the GPCR array as described above. However, the kit may further comprise reagents for generating a labelled target polynucleotide sample and/or a hybridisation buffer suitable for performing hybridisation between a biological material and the GPCR array.

EXAMPLES

The invention is further illustrated with reference to the following examples, which are not intended to be in any way limiting to the scope of the invention as claimed.

Example 1

Generation of Fragments for th GPCR Array (IC array)

Regions of opioid receptor polynucleotides were identified using a method as described in the following in general terms.

For each GPCR family, all members are identified, and the cDNA sequence of the members were downloaded from Genbank and the sequences of the Open Reading Frames (ORF) were compared to the other family members by clustal alignment (Higgins, D. G., Sharp, P. M., Gene 1988 Dec. 15; 73(1); 237-244). For each family member in turn, the alignments were performed to identify non-conserved regions of the member in question having as low a sequence identity to the other members of the family as possible (intrafamily). The level of intrafamily identity varies from family to family and from family member to family member, but in general it is initially attempted to identify regions having a level of identity of below 60%. However, for some family members this is not possible, and in such cases regions having the lowest existing sequence identity are used. Subsequently, the regions identified on the basis of a low intrafamily sequence identity are compared to the corresponding GPCR family members from other species to determine the level of sequence identity. The level of interspecies identity varies from family to family and from family member to family member, but in general it is initially attempted to obtain regions having a level of identity of above 70%. If this condition is not met, a different region is selected on the basis of the intrafamily identity as described above, and the interspecies identity of the new region is determined. This procedure is repeated a number of times to optimise the region selected with respect to partly intrafamily identity and partly interspecies identity.

The opioid receptors belong to the GPCR gene family and is currently classified into three groups μ, δ and κ. It has been suggested that further subtypes like μ1 and μ2 exists but these will not be considered here.

The cDNA sequences corresponding to the μ-opioid receptor (MOR), δ-opioid receptor (DOR) and κ-opioid receptor (KOR) from the rat (r) and human (h) species were downloaded from Genbank, and the sequences of the Open Reading Frames (ORF) were compared to each other by clustal alignment (see e.g. *Higgins DG and Sharp PM*, Gene 1988 **73** (1) 237-244).

Identity	GenBank No.
hMOR	NM_000914
rMOR	NM_013071
hDOR	NM_000911
rDOR	NM_012617
hKOR	XM_011716
rKOR	L22536

The alignments were performed to identify non-conserved regions of the GPCR member, regions having less than 60% sequence homology to the other GPCR members and at the same time having more than 70% homology between identical GPCR members from rat and human. Two fragments A and B were selected which fulfils these criteria.

Identity	Fragment A – coding sequence	Fragment B – coding sequence
hMOR	1-399	791-1203
rMOR	1-393	785-1197
hDOR	1-336	728-1119
rDOR	1-336	728-1119
hKOR	1-366	766-1143
rKOR	1-366	766-1143

Fragment A – percent identity (consensus length 406)

Gene	hMOR	rMOR	hDOR	rDOR	hKOR	rKOR
hMOR	X	80	48	45	55	51
rMOR		X	44	43	57	55
hDOR			X	70	53	50
rDOR				X	50	50
hKOR					X	72
RKOR						X

Fragment B – percent identity (consensus length 420)

Gene	hMOR	rMOR	hDOR	rDOR	hKOR	rKOR
hMOR	X	78	45	47	50	49
rMOR		X	49	51	51	53
hDOR			X	83	47	47
rDOR				X	49	47
hKOR					X	72
RKOR						X

Fragment A has a maximal of 57% identity between two intrafamily members (compare hKOR and rMOR) while the minimum interspecies identity is 70% (compare hDOR and rDOR). Thus the minimum difference in percentage intrafamily and interspecies identity is 13% which is sufficient to avoid cross hybridisation

between intrafamily members and a the same time get signal from interspecies members.

Fragment B has a maximal of 53% identity between two intrafamily members (compare rMOR and rKOR) while the minimum interspecies identity is 72% (compare hKOR and rKOR). Thus the minimum difference in percentage intrafamily and interspecies identity is 19% which is sufficient to avoid cross hybridisation between intrafamily members and a the same time get signal from interspecies members.

Example 2

Amplification of the determined regions

The two regions A (consensus sequence: 1-406) and B (consensus sequence: 801-1216) were amplified by PCR using following primers:

Region	Primer 1	Primer 2
hMOR-fragment A	5' ATG GAC AGC AGC GCT GCC C	5' TCC CAT TAG GTA ATT CAC AC
hMOR-fragment B	5' TCC GCA TGC TCT CTG GCT GC	5' TTA GGG CAA CGG AGC AGT TTC
hDOR-fragment A	5' ATG GAA CCG GCC CCC TCC	5' CTC CAT CAG GTA CTT GGC AC
hDOR-fragment B	5' TGC GCC TGC TGT CGG GCT C	5' TCA GGC GGC ACG GCC ACC
hKOR-fragment A	5' ATG GAC TCC CCG ATC CAG	5' ATT CAT CAA GTA GAC CGT AC
hKOR-fragment B	5' TCC GGC TCC TTT CTG GCT C	5' TCA TAC TGG TTT ATT CAT CC

Human whole brain mRNA was purchased from Clontech Laboratories Cat. No. 6516-1.

The cDNA was synthesized using the Omniscript RT Kit (Qiagen, 205111).

The amplification of the two regions were performed using:

1-2 µl	cDNA obtained above
250 µM	dNTP (27-2035-01, Amersham Pharmacia)
0.5 µM	of each PCR primers (see table above)
1.5 mM	$MgCl_2$ final concentration (Y02016, Gibco BRL)
1X	PCR buffer without $MgCl_2$ (Y02028, Gibco BRL)

2.5 U Taq polymerase (10966-026, Gibco BRL)

H_2O to 20 µl final volume

Using PCR conditions according to the supplier's manual.

The PCR generated fragments were separated using on a conventional agarose gel and cloned into the pCRII-TOPO vector according to the TOPO TA cloning kit (Invitrogen) and the nucleotide sequence was analysed using CEQ 2000 DNA Analysis System (Beckman Coulter, U.S.A.).

Example 3

Preparation of Master Glycerol Stocks

The glycerol stocks were prepared in 96 wells-trays (Corning Cat. No. cci3793) on a Biomek (Beckman Coulter, USA). 50 µl glycerol media was transferred into each well of a plane 96-well tray (Corning Cat. No. cci3793). 50 µl bacterial culture was transferred into each well of the plane 96-well tray and mixed with 2 x 100µl. A Storage Mat-I lid (Corning, Cat. No. 3094) was placed on each tray and the trays stored at -80°C.

Example 4

Preparation of plasmids for the GPCR Array

Ampicillin (100mg/ml) was added to Circlegrow medium (obtained from Bio101, Carlsbad, CA 92008, U.S.A.). Circlegrow medium/ampicillin was added to each well in a 4 x 2 ml 96 deep-well tray (Corning Cat. No. cci 3961). The glycerol stocks were added to each well. The tray was sealed with sealing sheet (Merck Eurolab A/S, Denmark), and incubated with shaking at 37°C for 16 hours prior to plasmid purification.

Example 5

Preparation of 3D-Link Amine-Binding array

The plasmids obtained in Example 4 were subjected to PCR using flanking primers. The resulting PCR product was spotted onto a 3D-Link Amine-Binding slide (array). The PCR reaction and spotting were carried out using standard methods as described e.g. in "Microarray Biochip Technology" by Mark Schena and "DNA Microarrays: A Practical Approach" by Mark Schena.

3D-Link Amine-Binding slide (array) was obtained from SurModics, Inc, Minneapolis, USA.

To have an orientation on the arrays visible after scanning of the slides the each of the arrays corners are marked by a double labelled Cy3 and Cy5 primer with a

5' end amino group (5'-Cy3/5) in a final concentration of 1pmol/µl to enable the possibility to place a grid on the scanned array.

After spotting, the slide is sealed with sealing tape and stored at -20°C until use.

3MM paper is pre-wetted with saturated NaCl solution. All slides are places in a slide box without a lid, and the slide boxes are placed in a plastic bag containing the NaCl saturated 3MM paper. The plastic bag is closed and the slides are incubated.

After incubation the slides are removed from the plastic bag and stored at room temperature.

The slides are placed in pre-heated blocking solution (0.1% SDS, SurModics Blocking Solution), incubated for 20 minutes at 50°C, and washed in redistilled H_2O. The slides are incubated in 4 x SSC, 0.1% SDS solution (50°C), washed at room temperate in redistilled H_2O. The slides are incubated in boiling redistilled H_2O for 2 minutes, and washed in redistilled H_2O at room temperature. The slides are incubated in pre-hybridisation buffer pre-heated to 50°C (50 ml 20 x SSC, 10 ml 100 x Denhardt solution, 2 ml 10% SDS, 4 ml salmon sperm DNA (10 mg/ml), 134 ml Redistilled H_2O) at 50°C for 30 minutes. The slides are washed in redistilled H_2O, and stored at room temperature in a dry and dark place until further use.

Example 6

Preparation of labelled samples and hybridisation

Whole human brain total RNA is obtained from Clontech Laboratories Cat. No. 64020-1). RNA is precipitated by centrifugation and the RNA pellet is washed in 70% ethanol. RNA is precipitated at 15,000 x g for 15 minutes. The supernatant is discarded and the pellet let to air-dry. The RNA concentration is adjusted to 1µg/µl with DEPC-H_2O. In 2 separate tubes 25 µl total RNA (1µg/µl) and 7 µl DEPC treated H_2O are added. 4 µl of oligo-dT (e.g. T25V primer) (1 µg/µl) are added to each tube. The tubes are incubated in a Thermal cycler at 65°C for 3 minutes.

Tube 1 is prepared by adding, 5 µl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM $MgCl_2$; 40mM DTT), 2 µl Cy3-dUTP (1mM, Amersham Pharmabiotech Cat. No. PA53022), 5 µl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents are mixed and 2 µl reverse transcriptase (100 U/µl) are added.

Tube 2 is prepared by adding, 5 µl 10 x cDNA Buffer (500mM Tris-HCl, pH 8.3; 800 mM KCl; 100 mM $MgCl_2$; 40mM DTT), 2 µl Cy5-dUTP (1mM, Amersham Pharmabiotech Cat. No. PA55022), 5 µl 10 x dNTP (5mM dATP; 5 mM dCTP; 5 mM dGTP; 5mM dTTP). The contents are mixed and 2 µl reverse transcriptase (100 U/µl) are added.

Tubes 1 and 2 are incubated at 42 °C for 60 minutes, and at 65°C for 15 minutes. The temperature is decreased to 42°C. Reverse transcriptase (100 U/µl) is added to each tube followed by incubation at 42°C for 60 minutes and at 65°C for 15 minutes. Precipitate with 3M Na-acetate and 96% ethanol. Wash each pellet in 80% ethanol.

Each pellet is resuspended in RNase Mix (10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA (pH 8.0), RNase A 100 mµ/ml) and incubated at 37°C for 60 minutes. 30 µl of sterile H_2O are added to each tube. Precipitation is accomplished using 3M Na-Acetate (pH 6.0) and ice-cold 96% ethanol. The pellets are washed in 80% ethanol, and resuspended in 15 µl hybridisation buffer (5 x SSC, 0.1% SDS, 100 µg/ml, blocking RNA).

The two fluorescents probes are mixed 1:1 in a PCR tube. This is the Probe-Mix.

The Probe-Mix is denatured at 100°C for 3 minutes, followed by a temperature decrease to 55°C for 30 seconds, where after the Probe-Mix is placed on ice. The Probe-Mix is added to the array slide, and the slides are placed in a box and inside the petri dish with the pre-wetted 3MM paper. The lid back is replaced onto the petri dish, and the petri dish is placed in a plastic bag.

The petri dish is incubated in a dark incubator at 65°C 12-16 hours.

The slides are washed in Washing Buffer I (2xSSC), and submerged in pre-warmed Washing Buffer II (2 x SSC, 0.1% SDS). The slides are pre-heated at 65°C for 1 hour in a volume of least 10 ml/slide to cover the slides, and incubated on an orbital shaker at 65°C for 10 minutes. The slides are washed in Washing Buffer III (0.2 x SSC) in a volume of least 10 ml/slide to cover the slides, and incubated on an orbital shaker at room temperature for 3 minutes. The slides are washed in Washing Buffer IV (0.1 x SSC), and in Washing Buffer V (0.5 x SSC). Washing with Washing Buffer V is repeated for additional 3 times, and all Washing Buffer V is removed by centrifuging at 800 rpm for 3 minutes.

Scanning of the slides and evaluation are performed using an Affymetrix 418 Scanner, Affymetrix 418 Scanner Software and ImaGene 4.0 software (BioDiscovery) according to the supplier's manual.

CLAIMS

1. A GPCR array comprising a multiplicity of individual GPCR polynucleotide spots stably associated with a surface of a solid support, wherein an individual GPCR polynucleotide spot comprises a GPCR polynucleotide composition comprising a non-conserved region of a GPCR polynucleotide family member, the spots representing at least two different regions of a GPCR polynucleotide family member.

2. The array according to claim 1, wherein the multiplicity of individual spots represents at least two different GPCR polynucleotide family members and/or at least two different GPCR families.

3. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has less than 90% intrafamily identity.

4. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has less than 85% intrafamily identity.

5. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has less than 80% intrafamily identity.

6. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has less than 75% intrafamily identity.

7. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has less than 50% intrafamily identity.

8. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has at least 50% interspecies identity.

9. The array according to claim 8, wherein the non-conserved regions of the GPCR family member has at least 60% interspecies identity.

10. The array according to claim 9, wherein the non-conserved regions of the GPCR family member has at least 65% interspecies identity.

11. The array according to claim 10, wherein the non-conserved regions of the GPCR family member has at least 70% interspecies identity.

12. The array according to claim 11, wherein the non-conserved regions of the GPCR family member has at least 75% interspecies identity.

13. The array according to claim 12, wherein the non-conserved regions of the GPCR family member has at least 80% interspecies identity.

14. The array according to any of the preceding claims, wherein the non-conserved regions of the GPCR polynucleotide family member has an average length of from about 3 to about 9,000 nucleotides, from about 3 to about 6,000, from about 3 to about 3,000, from about 200 to 750, from about 200 to 700, from about 200 to 500, from about 250 to 400 or from about 290 to 350.

15. The array according to any of the preceding claims, wherein said different regions of a GPCR polynucleotide family member are polynucleotide regions from the same polynucleotide strand and the regions are at least non-overlapping polynucleotide regions of the strand.

16. The array according to any of the preceding claims, wherein a GPCR polynucleotide family is polynucleotides encoding polypeptides capable of transducing a signal across biological membranes.

17. The array according to any of the preceding claims, wherein the GPCR polynucleotide family is selected from the group consisting of GPCR's belonging to any of the overall groups of Rhodopsin-like receptors, secretin-like, metabotropic glutamate/pheromone, Frizzled/smoothened and a large groups of unclassified GPCR's, which are all characterized by they seven transmembrane spanning regions and their ability to transduce a signal across a cellular membrane.

18. The array according to any of the preceding claims, wherein the polynucleotide composition comprises one or more of the same non-conserved region of a GPCR polynucleotide family member.

19. The array according to claim 18, wherein the polynucleotide composition comprises one or more of the same non-conserved region in the single stranded or double stranded form.

20. The array according to any of the preceding claims, wherein the non-conserved region of a GPCR polynucleotide family member is of DNA, RNA, cDNA, natural, synthetic, semi-synthetic origin or is a chemical analogous such as LNA and PNA.

21. The array according to any of the preceding claims, wherein the non-conserved region of a GPCR polynucleotide family member is obtained from one or more biological materials such as e.g. an organism, an organ, a tissue, a cell or a biological material produced by a cell culture.

22. The array according to claim 21, wherein the biological material is an organism, such as a microorganism, a plant, a fungus or an animal.

23. The array according to claim 22, wherein the biological material is an animal.

24. The array according to claim 23, wherein the animal is selected from the group consisting of humans, rats, mice, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, apes, monkeys, cats and sheep.

25. The array according to any of the preceding claims, wherein the solid support is made of a flexible or rigid material.

26. The array according to any of the preceding claims, wherein said array comprises from about 2 to about 100,000 such as, e.g. from about 2 to about 50,000, from about 10 to about 25,000, from about 100 to about 10,000, from about 100 to about 5,000, from about 100 to about 1,000, from about 400 to about 600 or about 500 GPCR polynucleotide spots, or at least 2 such as, e.g. at least 10, at least 25, at least 50, at least 100, at least 300, at least 400, at least 500 or at least 600 spots.

27. A method of preparing an array according to any of the preceding claims, said method comprising a) generating said non-conserved regions of GPCR polynucleotide family members, b) preparing a multiplicity of compositions each comprising at least a non-conserved region, and c) stably associating said compositions in individual spots on a surface of a solid support.

28. The method according to claim 27, wherein each said non-conserved region of a GPCR polynucleotide family member is produced by one or more primers specific for said non-conserved region.

29. A set of primers specific for non-conserved regions of GPCR polynucleotide family members, wherein the set of primers are used in the method according to any of the claims 27-28 for the production of an array according to any of the claims 1-26.

30. A method for the determination of a GPCR polynucleotide expression profile in a biological material, said method comprising a) obtaining a polynucleotide sample from the biological material, b) labelling said sample to obtain a labelled target polynucleotide sample, c) contacting at least one labelled target polynucleotide sample with an array according to any of the claims 1-25 under conditions which are sufficient to produce a hybridisation pattern, and d) detecting said hybridisation pattern to obtain the GPCR polynucleotide expression profile of the biological material.

31. A method for the determination of a difference in GPCR polynucleotide expression profiles from at least a first and a second different biological material, said method comprising obtaining a first GPCR expression profile of the first biological material according to the method of claim 29, obtaining a second GPCR expression profile of the second biological material according to the method of claim 29, comparing the first and the second GPCR expression profiles to identify any difference in the GPCR expression profiles between the first and the second GPCR expression profiles.

32. The method according to any of the claims 30-31, wherein the first and the second biological material is of the same kind of biological material.

33. The method according to the claims 30-32, wherein the first biological material is in a non-stressed state and the second biological material is in a stressed state.

34. The method according to the claim 32, wherein the stress directly or indirectly influence the GPCR expression profile of the first and/or the second biological material.

35. The method according to any of the claims 30-34, wherein the GPCR polynucleotide family is selected from the group consisting of GPCR's belonging to any of the overall groups of Rhodopsin-like receptors, secretin-like, metabotropic glutamate/pheromone, Frizzled/smoothened and a large groups of unclassified GPCR's, which are all characterized by they seven transmembrane spanning regions and their ability to transduce a signal across a cellular membrane.

36. The method according to claim 35, wherein the GPCR expression profile of the second biological sample is directly or indirectly related to a disease, chemical treatment, biological sample or parts in a biological sample treatment, environmental influences or other physiological or pathophysiological changes.

37. The method according to claim 36, wherein the chemical treatment is selected from the group consisting of naturally occurring chemical entities or synthetically derived chemical entities.

38. The method according to claim 36, wherein the disease is selected from the group consisting of asthma, cystic fibrosis, chronic obstructive pulmonary disease and rhinorrhea, convulsions, vascular spasms, coronary artery spasms, renal disorders, polycystic kidney disease, bladder spasms, urinary incontinence, bladder outflow obstruction, irritable bowel syndrome, gastrointestinal dysfunction, secretory diarrhoea, ischaemia, cerebral ischaemia, ischaemic hearth disease, angina pectoris, coronary hearth disease, traumatic brain injury, psychosis, anxiety, depression, dementia, memory and attention deficits, drug addiction and/or abuse, including cocaine or tobacco abuse, Parkinson's disease, Alzheimer's disease, dysmenorrhoea, narcolepsy, Reynaud's disease, intermittent claudication, Sjorgren's syndrome, migraine, arrhythmia, hypertension, absence seizures, myotonic muscle dystrophia, xerostomi, diabetes type II, hyperinsulinemia, premature labour, baldness, cancer, schizophrenia or psychosis; a variety of disorders associated with the neural

system, for example eating disorders, obsessive compulsive disorders, panic disorders, alcoholism, pain, memory deficits and anxiety including disorders such as pseudodementia or Ganser's syndrome, migraine pain, bulimia, obesity, pre-menstrual syndrome or late luteal phase syndrome, post-traumatic syndrome, memory loss, memory dysfunction, social phobia, attention deficit hyperactivity disorder, chronic fatigue syndrome, premature ejaculation, erectile difficulty, anorexia nervosa, disorders of sleep, autism, mutism, trichotillomania or mood syndrome; auto-immune diseases, e.g. Addison's disease, alopecia areata, Ankylosing spondylitis, haemolytic anaemia (anaemia haemolytica), pernicious anaemia (anaemia perniciosa), aphthae, aphthous stomatitis, arthritis, arteriosclerotic disorders, osteoarthritis, rheumatoid arthritis, aspermiogenese, asthma bronchiale, auto-immune asthma, auto-immune haemolysis, Becket's disease, Boeck's disease, inflammatory bowel disease, Burkett's lymphoma, Chron's disease, chorioiditis, colitis ulcerosa, Coeliac disease, cryoglobulinemia, dermatitis herpetiformis, dermatomyositis, insulin-dependent type I diabetes, juvenile diabetes, idiopathic diabetes insipidus, insulin-dependent diabetes mellisis, auto-immune demyelinating diseases, Dupuytren's contracture, encephalomyelitis, encephalomyelitis allergica, endophthalmia phacoanaphylactica, enteritis allergica, auto-immune enteropathy syndrome, erythema nodosum leprosum, idiopathic facial paralysis, chronic fatigue syndrome, febris rheumatica, glomerulo nephritis, Goodpasture's syndrome, Graves' disease, Hamman-Rich's disease, Hashimoto's disease, Hashimoto's thyroiditis, sudden hearing loss, sensoneural hearing loss, hepatitis chronica, Hodgkin's disease, haemoglobinuria paroxysmatica, hypogonadism, ileitis regionalis, iritis, leucopoenia, leukaemia, lupus erythematosus disseminatus, systemic lupus erythematosus, cutaneous lupus erythematosus, lymphogranuloma malignum, mononucleosis infectiosa, myasthenia gravis, traverse myelitis, primary idiopathic myxedema, nephrosis, ophthalmia symphatica, orchitis granulomatosa, pancreatitis, pemphigus, pemphigus vulgaris, polyarteritis nodosa, polyarthritis chronica primaria, polymyositis, polyradiculitis acuta, psoreasis, purpura, pyoderma gangrenosum, Quervain's thyreoiditis, Reiter's syndrome, sarcoidosis, ataxic sclerosis, progressive systemic sclerosis, scleritis, sclerodermia, multiple sclerosis, sclerosis disseminata, acquired spenic atrophy, infertility due to antispermatozoan antibodies, thrombocytopenia, idiopathic thrombocytopenia purpura, thymoma, acute anterior uveitis, vitiligo, AIDS, HIV, SCID and Epstein Barr virus associated diseases such as Sjorgren's syndrome, virus (AIDS or EBV) associated B cell lymphoma, parasitic diseases such as Lesihmania, and immune-suppressed

disease states such as viral infections following allograft transplantations, graft vs. Host syndrome, transplant rejection, or AIDS, cancers, chronic active hepatitis diabetes, toxic chock syndrome, food poisoning, and transplant rejection.

39. A method for identifying a therapeutic, prophylactic and/or toxic agent involved in a direct or indirect action on the GPCR expression profile in a biological material, said method comprises obtaining a first GPCR expression profile of a first biological material according to the method of claim 30, obtaining a second GPCR expression profile of a second biological material according to the method of claim 29, applying a test compound to the second biological material and obtaining a third GPCR expression profile thereof according to the method of claim 29, comparing the first, second and third GPCR expression profiles, and identifying any differences in the GPCR expression profiles so as to identify any biological response of the test compound on the GPCR expression profile.

40. The method according to claim 39 further comprising applying a test compound to the first biological material and obtaining a fourth GPCR expression profile thereof according to the method of claim 30, comparing the first, second, third and fourth GPCR expression profiles, and identifying any differences in the GPCR expression profiles so as to identify any biological response of the test compound on the GPCR expression profile.

41. The method according to claims 39 or 40, wherein the first and the second biological material is of the same kind of biological material.

42. The method according to the claims 38-41, wherein the first biological material is in a non-stressed state and the second biological material is in a stressed state.

43. The method according to the claim 42, wherein the stress directly or indirectly influence the GPCR expression profile of the first and/or the second biological material.

44. The method according to any of the claims 39-43, wherein the GPCR polynucleotide family is selected from the group consisting of GPCR's belonging to any of the overall groups of Rhodopsin-like receptors, secretin-like, metabotropic glutamate/pheromone, Frizzled/smoothened and a large groups of unclassified GPCR's, which are all characterized by they seven transmembrane

spanning regions and their ability to transduce a signal across a cellular membrane.

45. The method according to claim 44, wherein the GPCR expression profile of the second biological material is direct or indirect measure of a diseased state, a chemical pre-treatment, or environmental influences or other physiological or pathophysiological changes.

46. The method according to claim 45, wherein the disease is selected from the same group as defined in claim 38.

47. The method according to any of the claims 39-46, wherein the test compound is a chemical or biological derived compound such as compounds selected from the group consisting of therapeutic, prophylactic and/or toxic chemical entities, physiologically chemical entities, hormones, vitamins, nutrients, pesticides, fungicides, bateriocides and any other organic chemical entity.

48. The method according to any of the claims 30-47, wherein about 100 µg or less of total RNA of the biological material is used for hybridisation on each individual GPCR polynucleotide spot.

49. A diagnostic method to determine the differences of GPCR expression profiles between two biological materials; said method comprises obtaining a first GPCR expression profile of a first biological material according to the method of claim 30, obtaining a second GPCR expression profile of a second biological material according to the method of claim 30, comparing the first and second GPCR expression profile, and identifying any difference in the GPCR expression profiles.

50. The diagnostic method according to claim 49, wherein the difference between the GPCR expression profile of the first and the second GPCR expression profile is directly or indirectly influenced by a pathophysiological state or a disease such as diseases claimed in claim 38.

51. The diagnostic method according to claim 50, wherein the GPCR expression profile from more than two different biological materials are compared, such as biological materials, which are in different stages of a disease.

52. The method according to any of the claims 30–51, wherein the biological material is an organism, such as a microorganism, a plant, a fungus or an animal.

53. The method according to claim 52, wherein the animal is selected from the group consisting of humans, rats, mice, pigs, cows, horses, dogs, guinea pigs, ferrets, rabbits, apes, monkeys, cats and sheep.

54. A GPCR kit for use in a hybridisation assay, said kit comprising a GPCR array according to any of claims 1-26.

55. The GPCR kit according to claim 54, wherein said kit further comprises reagents for generating a labelled target polynucleotide sample.

56. The GPCR kit according to the claims 54-55, wherein said kit further comprises a hybridisation buffer.